

STIC-Biotech/ChemLib

126578

From: Borin, Michael  
Sent: Wednesday, July 07, 2004 11:51 AM  
To: STIC-Biotech/ChemLib  
Subject: Search request: 10/081108

Examiner: M.Borin  
AU: 1631  
Mailbox: 2C70  
Office: Remsen 2A55  
Tel.: 20713

CRFE

RE: 10/081108

Please search :

1. nucleic acids of SEQ ID NO: 1.
2. polypeptide SEQ ID NO: 2.

against the commercial protein and nucleic acid databases + interference search.

Thank you

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 7/7/04  
Date Completed: 7/7/04  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH: /  
NA Sequences: \_\_\_\_\_  
AA Sequences: /  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: cdp kasp  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

This Page Blank (uspto)



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 126578**

**TO: Michael Borin**  
**Location: REM-2A55&2C70**  
**Art Unit: 1631**  
**Friday, July 09, 2004**

**Case Serial Number: 10/081108**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Borin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**



```

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Thymic stromal-derived lymphopoietin, receptor (Fragment).
GN TSIPR
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
EX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koichiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR ENBL; AK010291; BAB26827.1; -.
DR MGD; MGI:1889506; Tslpr.
FT NON TER
SQ SEQUENCE 126 AA; 13321 MW; BD387D755A8D9DF6 CRC64;

Query Match 26.1%; Score 55.5; DB 11; Length 126;
Best Local Similarity 44.0%; Pred. No. 7.5;
Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

OY 21 KEESPVV---SWRLEPDGTALCFI 42
DB 68 EEEEDLIHTKAKRVEPDGTSLCTV 92

RESULT 14
Q82P22 PRELIMINARY; PRT; 317 AA.
AC Q82P22;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
GN SAV729.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
ON NCBI_TaxID=33903;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
EX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
EX MEDLINE=22608306; PubMed=12692562;

```

```

RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005023; BAC68439.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW Complete proteome.
SQ SEQUENCE 317 AA; 33592 MW; 9F6BF58EC301AD1 CRC64;

Query Match 26.1%; Score 55.5; DB 16; Length 317;
Best Local Similarity 46.3%; Pred. No. 21;
Matches 19; Conservative 6; Mismatches 11; Indels 5; Gaps 2;

OY 2 AARAVFLALSAQLQARLMKEESPVVSWRLEPED---GTAL 39
DB 108 AADLVGAALAPGLL--HLAQEAPGVSRRAEELEAGTAL 146

RESULT 15
Q9JMD5 PRELIMINARY; PRT; 359 AA.
AC Q9JMD5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytokine receptor delta1.
GN TSLPR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20197866; PubMed=10733486;
RA Fujio K., Nosaka T., Kojima T., Kawashima T., Yahata T.,
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Yamamoto K., Nishimura T.,
RA Kitamura T.;
RT "Molecular cloning of a novel type I cytokine receptor similar to the
RT common gamma chain.";
RL Blood 95:2204-2211(2000).
DR EMBL; AB031333; BAA92159.1; -.
DR PIR; JC7280; JC7280.
DR MGD; MGI:1889506; Tslpr.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003961; FN_III.
DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
KW Receptor.
SQ SEQUENCE 359 AA; 37792 MW; 37FC8FF507B0E42B CRC64;

Query Match 26.1%; Score 55.5; DB 11; Length 359;
Best Local Similarity 44.0%; Pred. No. 24;
Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

OY 21 KEESPVV---SWRLEPDGTALCFI 42
DB 301 EEEEDLIHTKAKRVEPDGTSLCTV 325

Search completed: July 7, 2004, 17:09:07
Job time : 48 secs

```

```

RA Ananides P.G., Scherer S.E., Li P.W., Hopkins R.A., Galle R.F.,
RA Sutton G.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieux E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Terrieria S., Fleischmann W.,
RA Fossier C., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Pittman K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Peterson G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
DR FLYBASE; FBgn0031107; HERC2.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0005012; P:ubiquitin cycle; IEA.
DR InterPro; IPR001199; Cyt_B5.
DR InterPro; IPR008979; Gal_bind like.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR000408; Reg Chr condens.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00632; HECT_1.
DR Pfam; PF00173; heme1_1.
DR Pfam; PF00415; RCC1_15.
DR PRINTS; PR00633; RCCNDNSATION.
DR SMART; SM00119; HECTC_1.
DR PROSITE; PS00237; HECT_1.
DR PROSITE; PS00626; RCC1_2_4.
DR PROSITE; PS00012; RCC1_3.
DR PROSITE; PS00030; UBA_1.
SQ SEQUENCE 4899 AA; 528653 MW; 2D10F544BDF71P4C CRC64;

Query Match 26.3%; Score 56; DB 5; Length 4899;
Best Local Similarity 43.8%; Pred. No. 3.6e+02;
Matches 14; Conservative 2; Mismatches 16; Indels 0; Gaps

QY 4 RAVFLALSQQLQARLMKEESPVTYSWRLEPED 35
DB 2528 RTVELAKQLSQAEIMPTPQIVQVWILEHPD 2559

RESULT 13
Q9CRJ6 PRELIMINARY; PRT; 126 AA.
ID Q9CRJ6
AC Q9CRJ6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

```

```

DE Brlr288 protein.
GN BUK1288.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
ON NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohata S., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR ENBL; AP005939; BAC46553.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti...; IEA.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00455; AMP_BINDING; 1__1; 1.
KW Complete proteome.
SQ SEQUENCE 619 AA; 65733 MW; C0A369E2FEAEED6 CRC64;

Query Match 27.0%; Score 57.5; DB 16; Length 619;
Best Local Similarity 53.3%; Pred.No. 22;
Matches 16; Conservative 2; Mismatches 11; Indels 1; Gaps

Qy 6 VFIALSAQLLQARLMKEESPVVSWRLEPED 35
Db 380 LFLALGLDLILQGYGMTETSPVSVN-TPED 408

RESULT 9
Q7U9Q3 PRELIMITARY; PRT; 340 AA.
AC Q7U9Q3;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Putative nucleotide sugar epimerase precursor (EC 5.1.3.-).
GN SYNW0200.
OS Synecococcus sp. (strain WHB102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
ON NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641;
RA Palenik B., Brahmsha B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
DR ENBL; BX569689; CAF06715.1; -.
KW Isomerase; Signal; Complete proteome.
FT SIGNAL 1 20 Potential.
SQ SEQUENCE 340 AA; 37286 MW; 6175034E966AA169 CRC64;

Query Match 26.8%; Score 57; DB 16; Length 340;
Best Local Similarity 32.4%; Pred.No. 14;
Matches 22; Conservative 6; Mismatches 14; Indels 26; Gaps

Qy 2 AARAVFVALSAQLL-----CARLMKEESPVV--SWRLEP----SD 35

```



```

DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008040; AAH08040.1; -.
DR GO; GO:0005829; C:cytosol; ISS.
DR GO; GO:0005624; C:membrane fraction; ISS.
DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.
DR GO; GO:000287; F:magnesium ion binding; ISS.
DR GO; GO:0007242; P:intracellular signaling cascade; ISS.
DR GO; GO:0046521; P:sphingoid catabolism; ISS.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGKC; 1.
DR ProDom; PD005043; DAGKC; 1.
DR SMART; SM00046; DAGKC; 1.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 384 AA; 42474 MW; F82999FF306113B0 CRC64;

Query Match 29.1%; Score 62; DB 4; Length 384;
Best Local Similarity 40.0%; Pred. No. 3.1;
Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

QY 3 ARAVFLALSQQLQARLMKEES-----PVVSWRLEPEDG 36
DQ 295 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 334
DI 295 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 334
DI 295 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 334

RESULT 3
Q96GK1 PRELIMINARY; PRT; 398 AA.
AC Q96GK1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009419; AAH09419.1; -.
DR GO; GO:0005829; C:cytosol; ISS.
DR GO; GO:0005624; C:membrane fraction; ISS.
DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.
DR GO; GO:000287; F:magnesium ion binding; ISS.
DR GO; GO:0007242; P:intracellular signaling cascade; ISS.
DR GO; GO:0046521; P:sphingoid catabolism; ISS.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGKC; 1.
DR ProDom; PD005043; DAGKC; 1.
DR SMART; SM00046; DAGKC; 1.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 398 AA; 43944 MW; 7F2C9B26C030E560 CRC64;

Query Match 29.1%; Score 62; DB 4; Length 398;
Best Local Similarity 40.0%; Pred. No. 3.2;
Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

QY 3 ARAVFLALSQQLQARLMKEES-----PVVSWRLEPEDG 36

```

```

DB 309 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 348
DI 309 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 348
DI 309 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 348

RESULT 4
Q8N632 PRELIMINARY; PRT; 470 AA.
AC Q8N632;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to sphingosine kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood, and Skin;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC030553; AAH30553.1; -.
DR EMBL; BC014439; AAH14439.1; -.
DR GO; GO:0005829; C:cytosol; ISS.
DR GO; GO:0005624; C:membrane fraction; ISS.
DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.
DR GO; GO:000287; F:magnesium ion binding; ISS.
DR GO; GO:0007242; P:intracellular signaling cascade; ISS.
DR GO; GO:0046521; P:sphingoid catabolism; ISS.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGKC; 1.
DR ProDom; PD005043; DAGKC; 1.
DR SMART; SM00046; DAGKC; 1.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; 1.
KW Kinase.
SQ SEQUENCE 470 AA; 51084 MW; 5172B93A38C7CC17 CRC64;

Query Match 29.1%; Score 62; DB 4; Length 470;
Best Local Similarity 40.0%; Pred. No. 3.8;
Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

QY 3 ARAVFLALSQQLQARLMKEES-----PVVSWRLEPEDG 36
DQ 381 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 420
DI 381 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 420
DI 381 SRAMLLRLFLAMEGRHMEYECPLYVYVVPVAFRLPEKDG 420

RESULT 5
Q87PB1 PRELIMINARY; PRT; 668 AA.
AC Q87PB1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tail-specific protease.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
DR EMBL; AP005078; BAC59869.1; -.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:04:24 ; Search time 39 seconds  
(without alignments)

347.879 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 MAARAVFLALSAQLQLQRLM.....SPVSVNRLEPEDGALCFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041

Total number of hits satisfying chosen parameters:

0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*\*

- 1: sp\_archaea:\*\*
- 2: sp\_bacteria:\*\*
- 3: sp\_fungi:\*\*
- 4: sp\_human:\*\*
- 5: sp\_invertebrate:\*\*
- 6: sp\_mammal:\*\*
- 7: sp\_mhc:\*\*
- 8: sp\_organelle:\*\*
- 9: sp\_phage:\*\*
- 10: sp\_plant:\*\*
- 11: sp\_rodent:\*\*
- 12: sp\_virus:\*\*
- 13: sp\_vertebrate:\*\*
- 14: sp\_unclassified:\*\*
- 15: sp\_rvirus:\*\*
- 16: sp\_bacteriap:\*\*
- 17: sp\_archaeap:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	29.1	290	4	Q9BTG7
2	62	29.1	384	4	Q96HV8
3	62	29.1	398	4	Q96GK1
4	62	29.1	470	4	Q9N632
5	62	29.1	668	16	Q87PB1
6	60	28.2	552	16	Q8ZOE6
7	58	27.2	173	5	Q9VYCL
8	57.5	27.0	619	16	Q89UX1
9	57	26.8	340	16	Q7U9Q3
10	56	26.3	404	16	Q7URB8
11	56	26.3	2019	5	Q9W0D3
12	56	26.3	4899	5	Q9VR91
13	55.5	26.1	126	11	Q9CRJ6
14	55.5	26.1	317	16	Q82P22
15	55.5	26.1	359	11	Q9JMD5
16	55.5	26.1	359	11	Q9JUH8

17	55.5	26.1	359	11	Q9JIE7
18	55.5	26.1	359	11	Q8CII9
19	55.5	26.1	370	11	Q9JIO7
20	55	25.8	666	16	Q8D9F9
21	55	25.8	1321	16	Q7UI85
22	55	25.8	3316	2	Q8VQF8
23	54.5	25.6	533	16	Q8ED79
24	54.5	25.6	877	5	Q8IAY9
25	54	25.4	525	10	Q8O664
26	54	25.4	552	16	Q8Z827
27	54	25.4	699	10	Q9FVX7
28	54	25.4	972	4	Q86YH7
29	53.5	25.1	308	16	Q827U3
30	53.5	25.1	664	16	Q7V8V8
31	53	24.9	362	5	Q9U483
32	53	24.9	552	16	Q8X6K5
33	53	24.9	552	16	Q8FJD6
34	53	24.9	552	16	Q83LR9
35	53	24.9	867	12	Q8AZL8
36	52.5	24.6	69	16	Q8D4B0
37	52.5	24.6	243	10	Q49572
38	52.5	24.6	539	16	Q822P6
39	52.5	24.6	924	5	Q9V783
40	52	24.4	346	10	Q9LFS7
41	52	24.4	445	10	Q9S2L2
42	52	24.4	804	3	Q9HGH7
43	52	24.4	945	16	Q26077
44	51.5	24.2	512	16	Q8RCI2
45	51.5	24.2	620	16	Q8A151

#### ALIGNMENTS

#### RESULT 1

Q9BTG7 PRELIMINARY; PRT; 290 AA.  
AC Q9BTG7  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DE 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Similar to sphingosine kinase 1 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue=Skin;  
RA Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC004112; AAH04112.1;  
DR GO; GO:0016301; P:kinase activity; IEA.  
KW Kinase.  
FT NON TER 1  
SQ SEQUENCE 290 AA; 31998 MW; A3B7A219DB52C520 CRC64;

Query Match 29.1%; Score 62; DB 4; Length 290;

Best Local Similarity 40.0%; Pred. No. 2.3; Indels 1;

Matches 16; Conservative 8; Mismatches 10; Gaps 6;

QY 3 ARAVFLALSAQLQLQRLMKEES-----PVVSWRLPEPDG 36

Db 201 SRAMLLRLFLAMEKGRHMEYECPLYLVVVFVFAFLPEPKDG 240

#### RESULT 2

Q96HV8 PRELIMINARY; PRT; 384 AA.  
AC Q96HV8

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 23:35:12 ; Search time 6472 seconds  
(without alignments)  
6723.796 Million cell updates/sec

Title: US-10-081-108-1  
Perfect score: 1004  
Sequence: 1 CGCATTAGGCTCTCCGG.....ATCTTTACACTAAAAGCC 1004

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:  
2: gb\_hg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vi:  
15: em\_ba:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_om:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:  
29: em\_vi:  
30: em\_htg\_hum:  
31: em\_htg\_inv:  
32: em\_htg\_other:  
33: em\_htg\_mus:  
34: em\_htg\_pln:  
35: em\_htg\_pro:  
36: em\_htg\_mam:  
37: em\_htg\_vrt:  
38: em\_sy:  
39: em\_htgo\_hum:  
40: em\_htgo\_mus:  
41: em\_htgo\_other:

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1004	100.0	1004	6	AR107991	AR107991 Sequence
2	1004	100.0	1004	6	AR167384	AR167384 Sequence
3	1004	100.0	1004	6	AR236412	AR236412 Sequence
4	1004	100.0	1004	6	AR411457	AR411457 Sequence
5	1004	100.0	1004	6	BD176601	BD176601 Method of
6	1004	100.0	1004	9	HSU19180	U19180 Human B mel
7	966	96.2	1032	6	AR104146	AR104146 Sequence
8	966	96.2	1032	6	I28470	I28470 Sequence 1
9	966	96.2	1032	6	I72215	I72215 Sequence 1
10	829.2	82.6	2154	9	AF527550	AF527550 Homo sapi
11	563.4	56.1	1068	9	AF527553	AF527553 Homo sapi
12	532.4	53.0	177896	2	AC064811	AC064811 Homo sapi
13	529.2	52.7	139063	2	AC133563	AC133563 Homo sapi
14	526	52.4	178300	2	AC145613	AC145613 Homo sapi
15	524.4	52.2	107701	2	AF155875	AF155875 Homo sapi
16	524.4	52.2	158712	2	AC074294	AC074294 Homo sapi
17	524.4	52.2	182078	2	AC140904	AC140904 Homo sapi
18	522.8	52.1	69684	9	AC006475	AC006475 Homo sapi
19	522.8	52.1	189396	2	AL161418	AL161418 Homo sapi
20	522.8	52.1	231566	2	AL158811	AL158811 Homo sapi
21	521.2	51.9	190277	9	AF254983	AF254983 Homo sapi
22	521.2	51.9	281116	9	HS21C001	HS21C001 Homo sapi
23	467.8	46.6	148606	9	AC104692	AC104692 Homo sapi
24	464.2	46.2	2267	9	AF527552	AF527552 Homo sapi
25	448	44.6	1104	9	AF527551	AF527551 Homo sapi
26	435.2	43.3	1840	9	AF339516	AF339516 Homo sapi
27	433.6	43.2	1840	9	AF339515	AF339515 Homo sapi
28	385	38.3	2071	9	AF527554	AF527554 Homo sapi
29	266	26.5	1891	9	AF339514	AF339514 Homo sapi
30	264.4	26.3	1891	9	AF218570	AF218570 Homo sapi
31	215.4	21.5	6318	9	AF499647	AF499647 Homo sapi
32	205.8	20.5	75657	9	HSB117C8	HSB117C8 Homo sapi
33	205.8	20.5	182078	2	AC140904	AC140904 Homo sapi
34	199.4	19.9	41185	9	HSB11C6	HSB11C6 Homo sapi
35	181.8	18.1	57728	6	E54646	E54646 Base sequen
36	181.8	18.1	129800	2	AL163539	AL163539 Homo sapi
37	181.8	18.1	142000	2	EX664725	EX664725 Homo sapi
38	181.8	18.1	186615	9	EX088717	EX088717 Human DNA
39	181.8	18.1	202891	9	EX088651	EX088651 Human DNA
40	179.4	17.9	42108	9	AC138036	AC138036 Homo sapi
41	179.4	17.9	120515	9	AC008443	AC008443 Homo sapi
42	179.4	17.9	236913	2	AL161615	AL161615 Homo sapi
43	179.2	17.8	11648	9	HS16C2	HS16C2 Homo sapi
44	179.2	17.8	190690	2	AL627234	AL627234 Homo sapi
45	177.6	17.7	119182	9	HSB62L20	AL050302 Homo sapi

ALIGNMENTS

RESULT 1  
LOCUS AR107991  
DEFINITION Sequence 1 from patent US 6110694.  
ACCESSION AR107991  
VERSION AR107991.1 GI:12823478  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1004)  
AUTHORS Boel, P., Wildmann, C., Boon-Falleur, T., van der Bruggen, P.,  
Coulie, P. and Renauld, J.-C.  
TITLE Methods for determining complexes of tumor rejection antigens and  
HLA-Cw\*160/molecules

Pred. No. is the number of results predicted by chance to have a



```
JOURNAL Patent: US 6110694-A 1 29-AUG-2000;
FEATURES      Location/Qualifiers
              1..1004
              /organism="unknown"
              /mol_type="unassigned DNA"
ORIGIN
```

Query Match	100.0%	Score 1004;	DB 6;	Length 1004;
Best Local Similarity	100.0%;	Prod. No. 3,7e-223;		
Matches 1004;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGCCAAATTTAGGGTCTCGGTATCTCCGGCTGAGCTGCTGTGTTCCCGGCTTAGAGACC	60	
Db	1	CGCCAAATTTAGGGTCTCGGTATCTCCGGCTGAGCTGCTGTGTTCCCGGCTTAGAGACC	60	
Qy	61	AGGAGAAAGGGGAGCTCGAGGCTGGAGCTGTATAACCGTGGCTGCTCACTCTGGATG	120	
Db	61	AGGAGAAAGGGGAGCTCGAGGCTGGAGCCTGTAAACCGTGGCTGCTCACTCTGGATG	120	
Qy	121	GTGGTGGCAACAGAGATGGCAGCCAGCTGGAGCTGTAAACCGTGGCTGCTCACTCTGGATG	180	
Db	121	GTGGTGGCAACAGAGATGGCAGCCAGCTGGAGCTGTAAAGAGGCGGCTGAGCGGTAGG	180	
Qy	181	AGTGGGGCTGGAGCAGTAAGATGCGGCGCAGAGCGGTTTTTCTGGCATTTGTCTGCCAGC	240	
Db	181	AGTGGGGCTGGAGCAGTAAGATGCGGCCAGAGCGGTTTTTCTGGCATTTGTCTGCCAGC	240	
Qy	241	TGCTCCAAAGCCAGGCTGATGAAGAGAGATCCCTGTGGTGGCTGAGCTGAGGTTGAGCGTG	300	
Db	241	TGCTCCAAAGCCAGGCTGATGAAGAGAGATCCCTGTGGTGGCTGAGGTTGAGGCGTG	300	
Qy	301	AAGACGSCACAGCTCTGTGCTTCACTTCTCAGGTTGTGGCAGCCACGGTGTAGGAGACG	360	
Db	301	AAGACGSCACAGCTCTGTGCTTCACTTCTCAGGTTGTGGCAGCCACGGTGTAGGAGACG	360	
Qy	361	GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTGACCGAGGATGGTCTC	420	
Db	361	GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTGACCGAGGATGGTCTC	420	
Qy	421	GATCTCTGACCTCGTGATCCGCGCGCTTGGCCCTCCAAAGTCCGAGATTTACAGCGAT	480	
Db	421	GATCTCTGACCTCGTGATCCGCGCGCTTGGCCCTCCAAAGTCCGAGATTTACAGCGAT	480	
Qy	481	GTGCATTTTGTAAAGCACTTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA	540	
Db	481	GTGCATTTTGTAAAGCACTTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA	540	
Qy	541	TGTATCAATATCTTGTGCTGACAGAGCGGCTCTTTTCAAGATTTCAATCACTCTCC	600	
Db	541	TGTATCAATATCTTGTGCTGACAGAGCGGCTCTTTTCAAGATTTCAATCACTCTCTCC	600	
Qy	601	TGCTTTGTCCAGAACACATTGACCAAGCTCTGAAAGATGTAAGTTTACGATAGAC	660	
Db	601	TGCTTTGTCCAGAACACATTGACCAAGCTCTGAAAGATGTAAGTTTACGATAGAC	660	
Qy	661	TTTTTAAACTTCAACCAATGTATTACTGAAATTAACAAATGTTGTAAATTCCTCGAGTGT	720	
Db	661	TTTTTAAACTTCAACCAATGTATTACTGAAATTAACAAATGTTGTAAATTCCTCGAGTGT	720	
Qy	721	TATTTCTACTGTATTAAGGTAATATACATAATCAATAAATCTGAGGATCATTTGCC	780	
Db	721	TATTTCTACTGTATTAAGGTAATATACATAATCAATAAATCTGAGGATCATTTGCC	780	
Qy	781	AGAGATTGTTGGGAGGGAAATGTTTATCAACGGTTTCAATGAAATTAATCCAAAAGTT	840	
Db	781	AGAGATTGTTGGGAGGGAAATGTTTATCAACGGTTTCAATGAAATTAATCCAAAAGTT	840	
Qy	841	ATTTTCCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATTCTTAATAACATTTTAAAA	900	
Db	841	ATTTTCCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATTCTTAATAACATTTTAAAA	900	
Qy	901	CCACTGTAGATGATGATAATAGGAGCTGTGAGTATTTCTGACATATCTATATAAATTA	960	

```

Db      901  CCACCTGTAGAAATGATGTAAATAGGAGCTGTCAGTATTTCTGCATATATCTATATAAAATTA 960
Qy      961  TTTAAAAGTCAATCAGTATTCAACATCTTTTACACATAAAAGCC 1004
          |||||||
Db      961  TTTAAAAGTCAATCAGTATTCAACATCTTTTACACATAAAAGCC 1004
          |||||||

RESULT 2
AR167384
LOCUS      AR167384                1004 bp      DNA      linear      PAT 17-DEC-2001
DEFINITION Sequence 20 from patent US 6287569.
ACCESSION  AR167384
VERSION     AR167384.1  GI:17903161
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1004)
AUTHORS     Kippes,T.J. and Wu,Y.
TITLE        Vaccines with enhanced intracellular processing
JOURNAL      Patent: US 6287569-A 20 11-SEP-2001;
FEATURES     Location/Qualifiers
              1..1004
               /organism="unknown"
               /mol_type="unassigned DNA"

ORIGIN

Query Match      100.0%; Score 1004; DB 6; Length 1004;
Best Local Similarity 100.0%; Pred. No. 3.7e-223;
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Query Match	100.0%;	Score 1004;	DB 6;	Length 1004;
Best Local Similarity	100.0%;	Pred. No. 3,7e-223;		
Matches 1004;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGCCAAATTTAGGGTCTCCGGTA	1	CGCCAAATTTAGGGTCTCCGGTATCTCCCGCTGAGCTGCTGTGTTC
DB	1	CGCCAAATTTAGGGTCTCCGGT	1	CGCCAAATTTAGGGTCTCCGGTATCTCCCGCTGAGCTGCTGTGTTC
QY	61	AGGAGAGGGGGAGCTGGAGGCT	61	AGGAGAGGGGGAGCTGGAGGCTGTAAACCGTGGCTCTCACTCTGGATG
DB	61	AGGAGAGGGGGAGCTGGAGGCT	61	AGGAGAGGGGGAGCTGGAGGCTGTAAACCGTGGCTCTCACTCTGGATG
QY	121	GTGCTGGCAACAGAGATGGCAG	121	GTGCTGGCAACAGAGATGGCAGCTGGAGTGTTAGAGGGCGGCTTGAGCGGTAGG
DB	121	GTGCTGGCAACAGAGATGGCAG	121	GTGCTGGCAACAGAGATGGCAGCTGGAGTGTTAGAGGGCGGCTTGAGCGGTAGG
QY	181	AGTGGGCTGGAGCAGTAAGATG	181	AGTGGGCTGGAGCAGTAAGATGGCGGCCAGAGCGGTTTTCTGGCAATTTCTG
DB	181	AGTGGGCTGGAGCAGTAAGATG	181	AGTGGGCTGGAGCAGTAAGATGGCGGCCAGAGCGGTTTTCTGGCAATTTCTG
QY	241	TGCTCCAAGCCAGGCTCATGA	241	TGCTCCAAGCCAGGCTCATGAAGGAGAGTCCCTGTGTGTGAGCTGGAGGTTG
DB	241	TGCTCCAAGCCAGGCTCATGA	241	TGCTCCAAGCCAGGCTCATGAAGGAGAGTCCCTGTGTGTGAGCTGGAGGTTG
QY	301	AAGACGGCAGACTGTGTGTTCA	301	AAGACGGCAGACTGTGTGTTCACTTCTGAGGTTGTGGCAGCAGGTGTAGGAGCG
DB	301	AAGACGGCAGACTGTGTGTTCA	301	AAGACGGCAGACTGTGTGTTCACTTCTGAGGTTGTGGCAGCAGGTGTAGGAGCG
QY	361	GCAGCTCAAAGGAGCAATAGG	361	GCAGCTCAAAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTACGCCAGGATG
DB	361	GCAGCTCAAAGGAGCAATAGG	361	GCAGCTCAAAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTACGCCAGGATG
QY	421	GATCTCTGACCTGTGATCGCC	421	GATCTCTGACCTGTGATCGCCCGGCTTGGCTTCCAAAGTGCAGAGATTACAGCGAT
DB	421	GATCTCTGACCTGTGATCGCC	421	GATCTCTGACCTGTGATCGCCCGGCTTGGCTTCCAAAGTGCAGAGATTACAGCGAT
QY	481	GTGCATTTTGTAAAGCACTTT	481	GTGCATTTTGTAAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA
DB	481	GTGCATTTTGTAAAGCACTTT	481	GTGCATTTTGTAAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA
QY	541	TGATCATTTATCTTGTGCTGC	541	TGATCATTTATCTTGTGCTGCAGGACCGGCTCTTTCAGGATTTTCAGTCACATCTTCC
DB	541	TGATCATTTATCTTGTGCTGC	541	TGATCATTTATCTTGTGCTGCAGGACCGGCTCTTTCAGGATTTTCAGTCACATCTTCC
QY	601	TGCTTTTCTCCAGAACACATTT	601	TGCTTTTCTCCAGAACACATTTGACCAAGCTCCTGAAAGATGTAAAGTTTACTACGCATAGAC



Db 601 TGCTTTGTCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGCATAGAC 660  
 Qy 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTTAAATTCCTCGAGTG 720  
 Db 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTTAAATTCCTCGAGTG 720  
 Qy 721 TATCTACTGTAATTTAAAGGTAATATACATAATCAATTAATCTGAGGGATCATTCGC 780  
 Db 721 TATCTACTGTAATTTAAAGGTAATATACATAATCAATTAATCTGAGGGATCATTCGC 780  
 Qy 781 AGAGATTGTTGGGGAGGGAATGTTATCAACCGGTTTCATTGAAATTAATCCAAAAAGTT 840  
 Db 781 AGAGATTGTTGGGGAGGGAATGTTATCAACCGGTTTCATTGAAATTAATCCAAAAAGTT 840  
 Qy 841 ATTTCCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATCTTAAACATTTTAAAAA 900  
 Db 841 ATTTCCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATCTTAAACATTTTAAAAA 900  
 Qy 901 CCACGTGAGATGATGTAATAGGAGCTGTGCAGTATTTCTGCATATATAATAATTA 960  
 Db 901 CCACGTGAGATGATGTAATAGGAGCTGTGCAGTATTTCTGCATATATAATAATTA 960  
 Qy 961 TTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004  
 Db 961 TTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004

RESULT 3  
 AR236412  
 LOCUS 1004 bp DNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 1 from patent US 6465184.  
 ACCESSION AR236412  
 VERSION AR236412.1 GI:27280387  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM  
 Boel, P., Wildmann, C., Boon-Falleur, T., van der Bruggen, P.,  
 Coulie, P., and Renaud, J.-C.  
 TITLE Isolated nucleic acid molecule encoding peptides which form  
 complexes with MHC molecules HLA-Cw\*1601 and uses thereof  
 JOURNAL Patent: US 6465184-A 1 15-OCT-2002;  
 FEATURES Location/Qualifiers  
 source 1..1004  
 /organism="unknown"  
 /mol\_type="genomic DNA"

Query Match 100.0%; Score 1004; DB 6; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-223;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CGCAATTTAGGCTCTCCGATATCTCCGCTGAGCTGCTGTTCCTCCGGCTTAGAGACC 60  
 Db 1 CGCAATTTAGGCTCTCCGATATCTCCGCTGAGCTGCTGTTCCTCCGGCTTAGAGACC 60  
 Qy 61 AGGAGAAGGGGAGCTGGAGGCTGTAAACCGTGGCTGCTCTCACTCTGGATG 120  
 Db 61 AGGAGAAGGGGAGCTGGAGGCTGTAAACCGTGGCTGCTCTCACTCTGGATG 120  
 Qy 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGATGTTAGAGGCGGCTTAGCGGTAG 180  
 Db 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGATGTTAGAGGCGGCTTAGCGGTAG 180  
 Qy 181 AGTGGGCTCGAGCAGTAAGATCGCGCCAGAGCGGTTTTCTGGCATTTCTGCCCCAGC 240  
 Db 181 AGTGGGCTCGAGCAGTAAGATCGCGCCAGAGCGGTTTTCTGGCATTTCTGCCCCAGC 240  
 Qy 241 TGCTCCAGCCAGCTGATGAAGAGAGTCCCTGTGTGTAGCTGGAGTTGGAGCTCG 300  
 Db 241 TGCTCCAGCCAGCTGATGAAGAGAGTCCCTGTGTGTAGCTGGAGTTGGAGCTCG 300

Qy 301 AAGACGGCAGAGCTCTGTGCTTCTATCTTCTGAGGTTGTGCGACCGCAGGTTGATGGACG 360  
 Db 301 AAGACGGCAGAGCTCTGTGCTTCTATCTTCTGAGGTTGTGCGACCGCAGGTTGATGGACG 360  
 Qy 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTGTCAGCCAGGATGGTCTC 420  
 Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTGTCAGCCAGGATGGTCTC 420  
 Qy 421 GATCTCTGACCTCGTGATTCGGCCCGCTTGGCCCTTCCAAAGTCCCGAGATTAACAGGAT 480  
 Db 421 GATCTCTGACCTCGTGATTCGGCCCGCTTGGCCCTTCCAAAGTCCCGAGATTAACAGGAT 480  
 Qy 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
 Db 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
 Qy 541 TGTATCATTAATCTTGTGCTGTCAGGAGCCGCTCTTTCAGGATTTCACTGACATCTTCC 600  
 Db 541 TGTATCATTAATCTTGTGCTGTCAGGAGCCGCTCTTTCAGGATTTCACTGACATCTTCC 600  
 Qy 601 TGCTTTGTCAGACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660  
 Db 601 TGCTTTGTCAGACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660  
 Qy 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTTGTAAATTCCTCGAGTG 720  
 Db 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTTGTAAATTCCTCGAGTG 720  
 Qy 721 TATCTACTGTAATTTAAAGGTAATATACATAATCAATTAATCTGAGGGATCATTCGC 780  
 Db 721 TATCTACTGTAATTTAAAGGTAATATACATAATCAATTAATCTGAGGGATCATTCGC 780  
 Qy 781 AGAGATTGTTGGGAGGGAATGTTATCAACCGTTCATTGAAATTAATCCAAAAAGTT 840  
 Db 781 AGAGATTGTTGGGAGGGAATGTTATCAACCGTTCATTGAAATTAATCCAAAAAGTT 840  
 Qy 841 ATTTCCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATCTTAAACATTTTAAAAA 900  
 Db 841 ATTTCCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATCTTAAACATTTTAAAAA 900  
 Qy 901 CCACGTGAGATGATGTAATAGGAGCTGTGCAGTATTTCTGCATATATAATAATTA 960  
 Db 901 CCACGTGAGATGATGTAATAGGAGCTGTGCAGTATTTCTGCATATATAATAATTA 960  
 Qy 961 TTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004  
 Db 961 TTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004

RESULT 4  
 AR411457  
 LOCUS 1004 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 1 from patent US 6638512.  
 ACCESSION AR411457  
 VERSION AR411457.1 GI:40163669  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM  
 Boel, P., Wildmann, C., Boon-Falleur, T., van der Bruggen, P.,  
 Coulie, P., and Renaud, J.-C.  
 TITLE Method for treating subjects by using BAGE tumor rejection antigen  
 precursors or tumor rejection antigens  
 JOURNAL Patent: US 6638512-A 1 28-OCT-2003;  
 FEATURES Location/Qualifiers  
 source 1..1004  
 /organism="unknown"  
 /mol\_type="genomic DNA"

Query Match 100.0%; Score 1004; DB 6; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-223;

Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 CGCCAAATTTAGGGTCTCCGGTATCTCCGGTGGAGCTGCTCTGTTCCGGCTTAGAGACC 60
Db	1 CGCCAAATTTAGGGTCTCCGGTATCTCCGGTGGAGCTGCTCTGTTCCGGCTTAGAGACC 60
Qy	61 AGGAGAGGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGCTCTCACTCTGGATG 120
Db	61 AGGAGAGGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGCTCTCACTCTGGATG 120
Qy	121 GTGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTTAGAGGGCGCGCTGAGCGGTAGG 180
Db	121 GTGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTTAGAGGGCGCGCTGAGCGGTAGG 180
Qy	181 AGTGGGCTGGAGCAGTAAGATGGCGCAGAGCGGTTTTTCTGGCAATGTCCTCCCGCAGC 240
Db	181 AGTGGGCTGGAGCAGTAAGATGGCGCAGAGCGGTTTTTCTGGCAATGTCCTCCCGCAGC 240
Qy	241 TGCTCCAAGCAGCGCTGATGAAGAGGAGTCCCCTGTGGTGGAGCTGGAGGTTGAGCGCTG 300
Db	241 TGCTCCAAGCAGCGCTGATGAAGAGGAGTCCCCTGTGGTGGAGCTGGAGGTTGAGCGCTG 300
Qy	301 AAGACGGCACAGCTCTGTGCTTCATCTTCTGAGTGTGGCAGCCACGGTGTAGAGAGCG 360
Db	301 AAGACGGCACAGCTCTGTGCTTCATCTTCTGAGTGTGGCAGCCACGGTGTAGAGAGCG 360
Qy	361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTCAAGGATGGTCTC 420
Db	361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTCAAGGATGGTCTC 420
Qy	421 GATCTCTGACCTCGTGATCGCGCGCTTGGCCCTCCAAAGTCCGAGATTACAGCGAT 480
Db	421 GATCTCTGACCTCGTGATCGCGCGCTTGGCCCTCCAAAGTCCGAGATTACAGCGAT 480
Qy	481 GTGCATTTGTGAACAATTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA 540
Db	481 GTGCATTTGTGAACAATTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA 540
Qy	541 TGTATCAATATCCTTTGTCTGCAGAGCGCGCTCCTTTCAAGTTCAGTTCACATCTTCC 600
Db	541 TGTATCAATATCCTTTGTCTGCAGAGCGCGCTCCTTTCAAGTTCAGTTCACATCTTCC 600
Qy	601 TGCTTTGTCCAGAACACATTGACCAAGCTCTGAAAGATGTAAAGTTTACTACGCATAGAC 660
Db	601 TGCTTTGTCCAGAACACATTGACCAAGCTCTGAAAGATGTAAAGTTTACTACGCATAGAC 660
Qy	661 TTTTAAACTTCAACCAATGTATTTACTGAAATAAACAAATGTTCTGAAATCCCTGAGTGT 720
Db	661 TTTTAAACTTCAACCAATGTATTTACTGAAATAAACAAATGTTCTGAAATCCCTGAGTGT 720
Qy	721 TATTTCTACTGTATTTAAAGGTAAATAACATAATCATTTAAATCTGAGGGATCATTTGCC 780
Db	721 TATTTCTACTGTATTTAAAGGTAAATAACATAATCATTTAAATCTGAGGGATCATTTGCC 780
Qy	781 AGAGATTGTTGGGAGGGAATGTATCAACGGTTTCAATTGAAATTTAAATCCAAAGATT 840
Db	781 AGAGATTGTTGGGAGGGAATGTATCAACGGTTTCAATTGAAATTTAAATCCAAAGATT 840
Qy	841 ATTTCTCAGAAAAATCAAAATAAAGTTTGCATGTTTTTATCTTTAAAAACATTTTAAAAA 900
Db	841 ATTTCTCAGAAAAATCAAAATAAAGTTTGCATGTTTTTATCTTTAAAAACATTTTAAAAA 900
Qy	901 CCACGTGAGAAATGATGTAATAGGAGCTGTGCAATTTCTTGACATATCTATATAAATTA 960
Db	901 CCACGTGAGAAATGATGTAATAGGAGCTGTGCAATTTCTTGACATATCTATATAAATTA 960
Qy	961 TTTAAAAAGTCAATCAGTATTCAACATCTTTTACATAAAAGCC 1004
Db	961 TTTAAAAAGTCAATCAGTATTCAACATCTTTTACATAAAAGCC 1004

RESULT 5  
BD176601

LOCUS	BD176601	1004 bp	DNA	linear	PAT 18-MAR-2000
DEFINITION	Method of molecular diagnosis of chronic myelogenous leukemia.				
ACCESSION	BD176601				
VERSION	BD176601.1	GI:29122311			
KEYWORDS	WO 02070747-A/3.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1004)				
AUTHORS	Mamo,H.				
TITLE	Method of molecular diagnosis of chronic myelogenous leukemia				
JOURNAL	Patent: WO 02070747-A 3 12-SEP-2002;				
	FUJISAWA PHARMACEUTICAL CO LTD,HIROYUKI MAMO				
COMMENT	OS Homo sapiens (human)				
	PN WO 02070747-A/3				
	PD 12-SEP-2002				
	PR 01-MAR-2002 WO 2002JP001901				
	PR 01-MAR-2001 JP OIP 056438				
	PI HIROYUKI MAMO				
	PC C12Q1/68,C12Q1/02,G01N33/15,C12N15/09,A61P35/02 CC				
	molecular diagnosis of chronic myelogenous leukemia FH				
	Method of				
	Location/Qualifiers				
FT	CDS	(201)..(332).			Key
FEATURES	Location/Qualifiers				
	1..1004				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	100.0%; Score 1004; DB 6; Length 1004;				
Best Local Similarity	100.0%; Pred. No. 3,7e-223;				
Matches 1004; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	CGCAATTATAGGCTCTCCGGTATCTCCGGCTAGCTGCTCTGTTCCCGGCTTAGAGACC	60		
Db	1	CGCAATTATAGGCTCTCCGGTATCTCCGGCTAGCTGCTCTGTTCCCGGCTTAGAGACC	60		
Qy	61	AGGAGAGGGGAGCTGGAGGCTGGAGCTGTAACACCGTGGCTCGTCTCACTCTGGATG	120		
Db	61	AGGAGAGGGGAGCTGGAGGCTGGAGCTGTAACACCGTGGCTCGTCTCACTCTGGATG	120		
Qy	121	GTGCTGCACACAGATGGCAGCCAGCTGGAGTGTTAGAGGCGGCTTGAGCGGTAGG	180		
Db	121	GTGCTGCACACAGATGGCAGCCAGCTGGAGTGTTAGAGGCGGCTTGAGCGGTAGG	180		
Qy	181	AGTGGGGCTGGACACTAGATGGCGCCAGAGCGGTTTTCTGGCATTTGCTGCCACG	240		
Db	181	AGTGGGGCTGGACACTAGATGGCGCCAGAGCGGTTTTCTGGCATTTGCTGCCACG	240		
Qy	241	TGCTCCAAAGCCAGGCTGATGAAGGAGAGTCCCTCTGTGTGAGCTGGAGGTTGGAGCCTG	300		
Db	241	TGCTCCAAAGCCAGGCTGATGAAGGAGAGTCCCTCTGTGTGAGCTGGAGGTTGGAGCCTG	300		
Qy	301	AAGACGCCACAGCTCTGTGCTTCACTTTGAGGTTGTGGCAGCCACGGTGTGGAGACG	360		
Db	301	AAGACGCCACAGCTCTGTGCTTCACTTTGAGGTTGTGGCAGCCACGGTGTGGAGACG	360		
Qy	361	GCAGCTCAACAGCAGCAATAGGAGGAGATGGAGTTTCACTGTGTGAGCCAGGATGGTCTC	420		
Db	361	GCAGCTCAACAGCAGCAATAGGAGGAGATGGAGTTTCACTGTGTGAGCCAGGATGGTCTC	420		
Qy	421	GATCTCTGACCTCGTGATCCGCGCCGCTTGGCCCTCCAAAGTGCAGAGATTACAGCGAT	480		
Db	421	GATCTCTGACCTCGTGATCCGCGCCGCTTGGCCCTCCAAAGTGCAGAGATTACAGCGAT	480		
Qy	481	GTCCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA	540		
Db	481	GTCCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA	540		
Qy	541	TGTATCATTTATCCTTGTGCTGACGAGCGGGCTCCTTTTTCAGGATTTTCAGTCACTTTC	600		

```
Db 541 TGATCATTTATCTGTGTGTCAGGAGCGCGCTCTTTCAGGATTTTCAGTCACATCTTCC 600
Qy 601 TGCTTTGTCCAGACACATGTCACCAAGCTCCTGAAGATGTAAGTTTACTAGCATAGAC 660
Db 601 TGCTTTGTCCAGAACACATGACCAAGCTCCTGAAGATGTAAGTTTACTAGCATAGAC 660
Qy 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAATAACAAATGTTGTAATTCCTCGAGTGT 720
Db 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAATAACAAATGTTGTAATTCCTCGAGTGT 720
Qy 721 TATTTACTTGTATTTAAAGGTAATTAATACATATCAATTTAAATCTGAGGGATCATTGC 780
Db 721 TATTTACTTGTATTTAAAGGTAATTAATACATATCAATTTAAATCTGAGGGATCATTGC 780
Qy 781 AGAGATGTTGGGGAGGGAATGTTATCAACGGTTCATGAAATTAATCAAAAAGTT 840
Db 781 AGAGATGTTGGGGAGGGAATGTTATCAACGGTTCATGAAATTAATCAAAAAGTT 840
Qy 841 ATTTCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATTTTAAACATTTTAAAA 900
Db 841 ATTTCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATTTTAAACATTTTAAAA 900
Qy 901 CCACTGTAGATCATGTAATAGGACTGTGCAGTATTTCTGACATATATATAAATTA 960
Db 901 CCACTGTAGATCATGTAATAGGACTGTGCAGTATTTCTGACATATATATAAATTA 960
Qy 961 TTTAAAGTCAATCAGTATTTCAACATCTTTTACACTTAAAAAGCC 1004
Db 961 TTTAAAGTCAATCAGTATTTCAACATCTTTTACACTTAAAAAGCC 1004

RESULT 6
HSU19180
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1004
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="patient M22"
/db_xref="taxon:9606"
/sex="female"
/cell_line="M22-MEL.43"
/tissue_type="melanoma"
/dev_stage="adult"
1. 1004
/gene="BAGE"
201. 332
/gene="BAGE"
/standard_name="B melanoma antigen"
/codon_start=1
/product="B melanoma antigen"
/protein_id="AAC50123.1"

gene
CDS
```

```
/db_xref="GI:726040"
/translation="MAARAVFLALSAQLLQARLMKESPVVSWRLEPEDGTALCFIR"
384. 485
/rpc_family="Alu"

repeat_region
ORIGIN
Query Match 100.0%; Score 1004; DB 9; Length 1004;
Best Local Similarity 100.0%; Pred. No. 3.7e-223;
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCAAATTTAGGCTCTCGGATATCTCCGGCTGAGCTCTGTTCGGGCTTAGAGACC 60
Db 1 CGCCAAATTTAGGCTCTCGGATATCTCCGGCTGAGCTCTGTTCGGGCTTAGAGACC 60
Qy 61 AGGAGAAAGGGGAGCTGGAGGCTGGAGCTGTAAACACCGTGGCTCTCACTCTGGATG 120
Db 61 AGGAGAAAGGGGAGCTGGAGGCTGGAGCTGTAAACACCGTGGCTCTCACTCTGGATG 120
Qy 121 GTGTGGCAACAGAGATGGCAGCGAGCTGGAGTGTGTAGAGGGCGGCTGAGCGGTAGG 180
Db 121 GTGTGGCAACAGAGATGGCAGCGAGCTGGAGTGTGTAGAGGGCGGCTGAGCGGTAGG 180
Qy 181 AGTGGGGCTGGAGCAGTAAGATGCGGCGCAGAGCGGTTTTTCTGGCATTGTCTGCCAGC 240
Db 181 AGTGGGGCTGGAGCAGTAAGATGCGGCGCAGAGCGGTTTTTCTGGCATTGTCTGCCAGC 240
Qy 241 TGCTCCAAAGCCAGGCTGATGAAGAGGAGTCCCCTGTGGTGAGCTGGAGGTTGAGCGCTG 300
Db 241 TGCTCCAAAGCCAGGCTGATGAAGAGGAGTCCCCTGTGGTGAGCTGGAGGTTGAGCGCTG 300
Qy 301 AAGAGCGCACAGCTCTGTGCTTTCATCTCTGAGGTTGTGGCAGCCAGGTCAGAGACG 360
Db 301 AAGAGCGCACAGCTCTGTGCTTTCATCTCTGAGGTTGTGGCAGCCAGGTCAGAGACG 360
Qy 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTGAGCAGGATGGTCTC 420
Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTGAGCAGGATGGTCTC 420
Qy 421 GATCTCTGACCTGTGATTCGCGCGCTTGGCTTCCAAAGTCCGAGATTCAGCGAT 480
Db 421 GATCTCTGACCTGTGATTCGCGCGCTTGGCTTCCAAAGTCCGAGATTCAGCGAT 480
Qy 481 GTGCATTTTGAAGCACCTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540
Db 481 GTGCATTTTGAAGCACCTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540
Qy 541 TGTATCATATCTTGTGCTGCTGAGAGCGGCTCTTTTCAAGATTCAGTCACATCTTCC 600
Db 541 TGTATCATATCTTGTGCTGCTGAGAGCGGCTCTTTTCAAGATTCAGTCACATCTTCC 600
Qy 601 TGCTTTGTCCAGAACACATTCACCAAGCTCCTGAAAGATGTAAGTTTACTACGATAGAC 660
Db 601 TGCTTTGTCCAGAACACATTCACCAAGCTCCTGAAAGATGTAAGTTTACTACGATAGAC 660
Qy 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAATAACAAATGTTGTAATTCCTCGAGTGT 720
Db 661 TTTTAAACTTCAACCAATGTAATTTACTGAAAATAACAAATGTTGTAATTCCTCGAGTGT 720
Qy 721 TATTTACTTGTATTTAAAGGTAATTAATACATATCAATTTAAATCTGAGGGATCATTGC 780
Db 721 TATTTACTTGTATTTAAAGGTAATTAATACATATCAATTTAAATCTGAGGGATCATTGC 780
Qy 781 AGAGATGTTGGGGAGGGAATGTTATCAACGGTTCATTTGAAATTAATCAAAAAGTT 840
Db 781 AGAGATGTTGGGGAGGGAATGTTATCAACGGTTCATTTGAAATTAATCAAAAAGTT 840
Qy 841 ATTTCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATTTTAAACATTTTAAAA 900
Db 841 ATTTCTCAGAAAAATCAATAAAGTTTGCATGTTTTTATTTTAAACATTTTAAAA 900
Qy 901 CCACTGTAGATCATGTAATAGGACTGTGCAGTATTTCTGACATATATATAAATTA 960
Db 901 CCACTGTAGATCATGTAATAGGACTGTGCAGTATTTCTGACATATATATAAATTA 960
```

```

QY 961 TTAARAGTCAATCAGTATTCAACATCTTTTACACTAAAAGCC 1004
Db 961 TTAARAGTCAATCAGTATTCAACATCTTTTACACTAAAAGCC 1004

RESULT 7
AR104146
LOCUS AR104146 1032 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6093540.
ACCESSION AR104146
VERSION AR104146.1 GI:12816854
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1032)
AUTHORS van der Bruggen,P., Boon-Falleur,T., Coullie,P. and Renauld,J.-C.
TITLE Method for diagnosing a disorder characterized by expression of a
BAGE tumor rejection antigen precursor
JOURNAL Patent: US 6093540-A 1 25-JUL-2000;
FEATURES
source Location/Qualifiers
1..1032
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 96.2%; Score 966; DB 6; Length 1032;
Best Local Similarity 97.3%; Pred. No. 2.6e-214;
Matches 1004; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 CGCCAAATTAGGGTCTCCGGTATCTCCCGTGAAGTCTCTGTCCTCCGGCTTAGAGACC 60
Db 1 CGCCAAATTAGGGTCTCCGGTATCTCCCGTGAAGTCTCTGTCCTCCGGCTTAGAGACC 60

QY 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGCTCGTCTCACTCTCGATG 120
Db 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGCTCGTCTCACTCTCGATG 120

QY 121 GTGTGTGCAACAGAGATGGCAGCGCAGCTGAGTGTGTAGGAGCGCGCTGAGCGTAGG 180
Db 121 GTGTGTGCAACAGAGATGGCAGCGCAGCTGAGTGTGTAGGAGCGCGCTGAGCGTAGG 180

QY 181 AGTGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTGTCTGCCACG 240
Db 181 AGTGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTGTCTGCCACG 240

QY 241 TGCTCCAAAGCCAGGCTGATGAAGGAGGAGTCCCTGTGTGTGAGCTGGAGGTTGGAGCCTG 300
Db 241 TGCTCCAAAGCCAGGCTGATGAAGGAGGAGTCCCTGTGTGTGAGCTGGAGGTTGGAGCCTG 300

QY 301 AAGAGCGCACAGCTCTGTGCTTCACTTCTGAGGTTGTGGCAGCCACGGTGTAGGAGCG 360
Db 301 AAGAGCGCACAGCTCTGTGCTTCACTTCTGAGGTTGTGGCAGCCACGGTGTAGGAGCG 360

QY 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTCTGACGAGGATGTCTC 420
Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTCTGACGAGGATGTCTC 420

QY 421 GATCTCTGACCTGTGATCCGCCCGCTTTGGCGCTTCCAAAGTCCGAGATTACAGCGAT 480
Db 421 GATCTCTGACCTGTGATCCGCCCGCTTTGGCGCTTCCAAAGTCCGAGATTACAGCGAT 480

QY 481 GTGCATTTTGTAGACCTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCAGA 540
Db 481 GTGCATTTTGTAGACCTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCAGA 540

QY 541 TGATCATATTCTTGTGCTGACGAGCGCGCTCTTTTCAAGGATTTCACTCACATCTCC 600
Db 541 TGATCATATTCTTGTGCTGACGAGCGCGCTCTTTTCAAGGATTTCACTCACATCTCC 600

QY 601 TGCTTTGTCCAGACACATTGACCAAGCTCTCAAGAGATGTAGTTTACTACGATAGAC 660

```

```

Db 601 TGCTTTGTCCAGAACACATTGACCAAGCTCTCTGAAGATGTAAAGTTTACTACGCATAGAC 660
QY 661 TTTTAAACTTCAACCAATGTATTTTACTGAAAATAACAAATGTTGTAAATTTCCCTGAGTGT 720
Db 661 TTTTAAACTTCAACCAATGTATTTTACTGAAAATAACAAATGTTGTAAATTTCCCTGAGTGT 720
QY 721 TATTCTACTTGTATTTAAAGGTAATAACATAATCAATAATCTTAAATCTGAGGGATCATTTGCC 780
Db 721 TATTCTACTTGTATTTAAAGGTAATAACATAATCAATAATCTTAAATCTGAGGGATCATTTGCC 780
QY 781 AGAGATTGTTGGGGAGGG-----AAATGTTATCAAG 812
Db 781 AGAGATTGTTGGGGAGGGAAATGTTATCAACGGTTTCATTGAAATTTAAATGTTATCAAG 840
QY 813 GTTTCATTGAAATTTAAATCCAAAAGTTATTTCCTCAGAAAATCAAAATAAAGTTTCAT 872
Db 841 GTTTCATTGAAATTTAAATCCAAAAGTTATTTCCTCAGAAAATCAAAATAAAGTTTCAT 900
QY 873 GTTTTTTATTTTAAACATTTTAAACCACTGTAGAAATGATAATAGGAGCTGTGC 932
Db 901 GTTTTTTATTTTAAACATTTTAAACCACTGTAGAAATGATAATAGGAGCTGTGC 960
QY 933 AGTATTTCTGACATATACATAATTAATAATTTAAATAAGTCAATCACTTCAACATCTTTTA 992
Db 961 AGTATTTCTGACATATACATAATTAATAATTTAAATAAGTCAATCACTTCAACATCTTTTA 1020
QY 993 CACTAAAAGGCC 1004
Db 1021 CACTAAAAGGCC 1032

RESULT 8
128470
LOCUS 128470 1032 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 1 from patent US 5571711.
ACCESSION 128470
VERSION 128470.1 GI:1819246
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1032)
AUTHORS van der Bruggen,P., Boon-Falleur,T., Coullie,P. and Renauld,J.-C.
TITLE Isolated nucleic acid molecules coding for BAGE tumor rejection
antigen precursors
JOURNAL Patent: US 5571711-A 1 05-NOV-1996;
FEATURES
source Location/Qualifiers
1..1032
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 96.2%; Score 966; DB 6; Length 1032;
Best Local Similarity 97.3%; Pred. No. 2.6e-214;
Matches 1004; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 CGCCAAATTAGGGTCTCCGGTATCTCCCGTGAAGTCTCTGTCCTCCGGCTTAGAGACC 60
Db 1 CGCCAAATTAGGGTCTCCGGTATCTCCCGTGAAGTCTCTGTCCTCCGGCTTAGAGACC 60

QY 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGCTCGTCTCACTCTCGATG 120
Db 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGCTCGTCTCACTCTCGATG 120

QY 121 GTGTGTGCAACAGAGATGGCAGCGCAGCTGAGTGTGTAGGAGCGCGCTGAGCGTAGG 180
Db 121 GTGTGTGCAACAGAGATGGCAGCGCAGCTGAGTGTGTAGGAGCGCGCTGAGCGTAGG 180

QY 181 AGTGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTGTCTGCCACG 240
Db 181 AGTGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTGTCTGCCACG 240

QY 241 TGCTCCAAAGCCAGGCTGATGAAGGAGGAGTCCCTGTGTGTGAGCTGGAGGTTGGAGCCTG 300

```

```
Db 241 TGCTCCAAAGCCAGCTGATGAAGAGAGTCCCTGTGGTGAAGTGGAGCTG 300
Qy 301 AAGAGGCGACAGCTCTGTGCTTCACTTCTGAGGTTGGCAGCCACGCGTATGAGAGC 360
Db 301 AAGAGGCGACAGCTCTGTGCTTCACTTCTGAGGTTGGCAGCCACGCGTATGAGAGC 360
Qy 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACCTGTCTCAGCCAGGATGCTCTC 420
Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACCTGTCTCAGCCAGGATGCTCTC 420
Qy 421 GATCTCTGACCTGTGATCGGCGGCTTGGGCTTCCAAAGTGGCCAGATTAACGCGAT 480
Db 421 GATCTCTGACCTGTGATCGGCGGCTTGGGCTTCCAAAGTGGCCAGATTAACGCGAT 480
Qy 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGACCCAGA 540
Db 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGACCCAGA 540
Qy 541 TGTATCATTTATCTTGTGCTGCAGGAGCCGCTTTCAGAGATTTCAATCATCTTCC 600
Db 541 TGTATCATTTATCTTGTGCTGCAGGAGCCGCTTTCAGAGATTTCAATCATCTTCC 600
Qy 601 TGCTTTGTCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660
Db 601 TGCTTTGTCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660
Qy 661 TTTTAAACTTCAACCAATGATTTACTGAAATTAACAAATGTTTAAATTTCCCTGAGTGT 720
Db 661 TTTTAAACTTCAACCAATGATTTACTGAAATTAACAAATGTTTAAATTTCCCTGAGTGT 720
Qy 721 TATCTACTTGTATTAAGAGTAAATATAATCAATCAATTAATACTGAGGATCAATTGCC 780
Db 721 TATCTACTTGTATTAAGAGTAAATATAATCAATCAATTAATACTGAGGATCAATTGCC 780
Qy 781 AGAGATTGTTGGGAGGCG-----AAATGTTATCAAG 812
Db 781 AGAGATTGTTGGGAGGAGAAATGTTATCAAGGTTTCAATGAAATTAATGTTATCAAG 840
Qy 813 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 872
Db 841 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 900
Qy 873 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 932
Db 901 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 960
Qy 933 AGTATTTCTGACATATACTATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTA 992
Db 961 AGTATTTCTGACATATACTATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTA 1020
Qy 993 CACTAAAGGCC 1004
Db 1021 CACTAAAGGCC 1032
```

```
RESULT 9
LOCUS 172215
DEFINITION Sequence 1 from patent US 5683886.
ACCESSION 172215
VERSION 172215.1
KEYWORDS GI:3008354
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS van der Bruggen,P. and Boon-Falleur,T.
TITLE Tumor rejection antigens which correspond to amino acid sequences in tumor rejection antigen precursor bage, and uses thereof
JOURNAL Patent: US 5683886-A 1 04-NOV-1997;
FEATURES Location/Qualifiers
source 1..1032
```

## ORIGIN

```
Query Match 96.2%; Score 966; DB 6; Length 1032;
Best Local Similarity 97.3%; Pred. No. 2,6e-214;
Matches 1004; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

Qy 1 CCCCAATTTAGGCTTCCGGTATCTCCCGCTGAGCTGCTCTGTGTTCCCGCTTAGAGACC 60
Db 1 CCCCAATTTAGGCTTCCGGTATCTCCCGCTGAGCTGCTCTGTGTTCCCGCTTAGAGACC 60
Qy 61 AGGAGAAAGGGGAGCTGGAGGCTGTAACACCGTGGCTGCTCACTCTGGAATG 120
Db 61 AGGAGAAAGGGGAGCTGGAGGCTGTAACACCGTGGCTGCTCACTCTGGAATG 120
Qy 121 GTGGTGGCAACAGAGATGGCAGCGAGCTGGAGTGTTAGGAGGGCGGCTGAGCGGTAGG 180
Db 121 GTGGTGGCAACAGAGATGGCAGCGAGCTGGAGTGTTAGGAGGGCGGCTGAGCGGTAGG 180
Qy 181 AGTGGGGCTGGAGCAGTAAAGATGGCGGCAGAGGGGTTTTTCTGGCATGTCTGCCACG 240
Db 181 AGTGGGGCTGGAGCAGTAAAGATGGCGGCAGAGGGGTTTTTCTGGCATGTCTGCCACG 240
Qy 241 TGCTCCAAAGCCAGGCTGATGAAGGAGGAGTCCCTGTGCTGAGCTGGAGGTTGAGCGCTG 300
Db 241 TGCTCCAAAGCCAGGCTGATGAAGGAGGAGTCCCTGTGCTGAGCTGGAGGTTGAGCGCTG 300
Qy 301 AAGAGCGGACAGCTCTGTGCTTCACTTCTGAGGTTGTGGCAGCCACGCTGATGGAGACG 360
Db 301 AAGAGCGGACAGCTCTGTGCTTCACTTCTGAGGTTGTGGCAGCCACGCTGATGGAGACG 360
Qy 361 CGAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTCCTGTGTCAGCCAGATGGTCTC 420
Db 361 CGAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTCCTGTGTCAGCCAGATGGTCTC 420
Qy 421 GATCTCTGACCTGTGATTCGCGCGGCTTGGCGCTTCCAAAGTCCGAGATTAACGCGAT 480
Db 421 GATCTCTGACCTGTGATTCGCGCGGCTTGGCGCTTCCAAAGTCCGAGATTAACGCGAT 480
Qy 481 GTGCATTTTGAAGCAGCTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCAGA 540
Db 481 GTGCATTTTGAAGCAGCTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCAGA 540
Qy 541 TGTATCATTTATCTTGTGCTGCAGGAGCGGCTTCTTTCAGGATTTCACTCAGTACATCTTCC 600
Db 541 TGTATCATTTATCTTGTGCTGCAGGAGCGGCTTCTTTCAGGATTTCACTCAGTACATCTTCC 600
Qy 601 TGCTTTGTCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660
Db 601 TGCTTTGTCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660
Qy 661 TTTTAAACTTCAACCAATGATTTACTGAAATTAACAAATGTTTAAATTTCCCTGAGTGT 720
Db 661 TTTTAAACTTCAACCAATGATTTACTGAAATTAACAAATGTTTAAATTTCCCTGAGTGT 720
Qy 721 TATCTACTTGTATTAAGAGTAAATATAATCAATCAATTAATACTGAGGATCAATTGCC 780
Db 721 TATCTACTTGTATTAAGAGTAAATATAATCAATCAATTAATACTGAGGATCAATTGCC 780
Qy 781 AGAGATTGTTGGGAGGCG-----AAATGTTATCAAG 812
Db 781 AGAGATTGTTGGGAGGAGAAATGTTATCAAGGTTTCAATGAAATTAATGTTATCAAG 840
Qy 813 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 872
Db 841 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 900
Qy 873 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 932
Db 901 GTTTCATTTGAAATTAATCCAAAGTATTTCTCAGAAAAATCAAAATTAAGTTTGCAT 960
Qy 933 AGTATTTCTGACATATACTATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTA 992
Db 961 AGTATTTCTGACATATACTATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTA 992
Qy 993 CACTAAAGGCC 1004
Db 1021 CACTAAAGGCC 1032
```

/organism="unknown"  
/mol\_type="unassigned DNA"

```

Db      961 AGTATTTCTGACATATACATAATAAATTATTAAGTCAATCAGTATTCAACATCTTTA 1020
Qy      993 CACTAAAAGCC 1004
Db      1021 CACTAAAAGCC 1032

RESULT 10
AF527550      2154 bp mRNA linear PRI 31-JAN-2003
LOCUS      Homo sapiens B melanoma antigen variant b (BAGE1) mRNA, complete
DEFINITION      cds, alternatively spliced.
ACCESSION      AF527550
VERSION      AF527550.1 GI:28173029
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 2154)
AUTHORS      Ruault,M., Van Der Bruggen,P., Brun,M.E., Boyle,S., Roizes,G. and
Sario,A.D.
TITLE      New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric
regions of human chromosomes 13 and 21 have a cancer/testis
expression profile
JOURNAL      Eur. J. Hum. Genet. 10 (12), 833-840 (2002)
MEDLINE      22349465
PubMed      12461691
REFERENCE      2 (bases 1 to 2154)
AUTHORS      Ruault,M., Ventura,M., Galtier,N., Brun,M.-E., Archidiacono,N.,
Roizes,G. and De Sario,A.
TITLE      BAGE, a gene family generated by juxtacentromeric reshuffling in
the Hominidae lineage, is under selective pressure
JOURNAL      Genomics (2003) In press
REFERENCE      3 (bases 1 to 2154)
AUTHORS      Ruault,M., van der Bruggen,P., Brun,M.-E., Boyle,S., Roizes,G. and
De Sario,A.
TITLE      Direct Submission
JOURNAL      Submitted (08-JUL-2002) Institut de Genetique Humaine, CNRS UPR
1142, 141, rue de la Cardonille, Montpellier 34396, France
FEATURES
source      Location/Qualifiers
1..2154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="13"
1..2154
/gene="BAGE1"
201..332
/gene="BAGE1"
/notes="tumor antigen; alternatively spliced"
/codon_start=1
/product="B melanoma antigen variant b"
/protein_id="AAO32670.1"
/db_xref="GI:28173030"
/translation="MARAVFLSALQQLQALMKKEESPVWSRLEPEDGALCFIF"

Query Match      82.6%; Score 829.2; DB 9; Length 2154;
Best Local Similarity 92.1%; Pred. No. 2e-182;
Matches 925; Conservative 0; Mismatches 8; Indels 71; Gaps 2;

Qy      1 CGCCAAATTAGGCTCCGGTATCTCCGGCTGAGCTCTGTGTCGGGCTTAGAGACC 60
Db      1 CGCCAAATTAGGCTCCGGTATCTCCGGCTGAGCTCTGTGTCGGGCTTAGAGACC 60

Qy      61 AGGAGAGGGGAGCTGGAGCTGTAAACACCGGTGGCTCGTCTCACTCTGGATG 120
Db      61 AGGAGAGGGGAGCTGGAGCTGTAAACACCGGTGGCTCGTCTCACTCTGGATG 120

Qy      121 GTGTTGGCAACAGATGGCAGCGCAGTGGTGTAGGAGCGCGCTGAGCGGTAGG 180

```

```

Db      121 GTGTTGGCAACAGATGGCAGCGCAGTGGAGTGTTAGAGGGCGGCTTGAGCGGTAGG 180
Qy      181 AGTGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTCTGCCCCAGC 240
Db      181 AGTGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTCTGCCCCAGC 240
Qy      241 TGTCTCAAAGCCAGGCTGATGAAGGAGGAGTCCCCTGTGTGAGCTGGAGGTTGGAGCCTG 300
Db      241 TGTCTCAAAGCCAGGCTGATGAAGGAGGAGTCCCCTGTGTGAGCTGGAGGTTGGAGCCTG 300
Qy      301 AAGAGGCGACAGCTCTGTGCTTCACTTTCTGAGGTTGTGGCAGCCACGCGTATGGAGACG 360
Db      301 AAGAGGCGACAGCTCTGTGCTTCACTTTCTGAGGTTGTGGCAGCCACGCGTATGGAGACG 360
Qy      361 GCAGCTCAACAGGAGCAATAGGAGGAGTGGAGTTTTCATCTGTGTCAAGAGGATGGTCTC 420
Db      361 GCAGCTCAACAGGAGCAATAGGAGG-----GTAAACCGTGGAGGCCA----- 401
Qy      421 GATCTCTGACCTCGTGTATCGGCCCGCTTGGCCCTTCCAAAGTCCGAGATTACAGCGAT 480
Db      402 -----AGTGGCAT 409
Qy      481 GTGCATTTTGTAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA 540
Db      410 GTGCATTTTGTAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCCAGA 469
Qy      541 TGTATCATTAATCTTGTGTGTCGAGGAGCGGCTCTTTTCAGGATTTTCAGTCACATCTTCC 600
Db      470 TGTATCATTAATCTTGTGTGTCGAGGAGCGGCTCTTTTCAGGATTTTCAGTCACATCTTCC 529
Qy      601 TGCTTTGTCCAGAACACATTGACCAAGCTCTCTGAAGATGTAACTTTTACTACGCATAGAC 660
Db      530 TGCTTTGTCCAGAACACATTGACCAAGCTCTCTGAAGATGTAACTTTTACTACGCATAGAC 589
Qy      661 TTTTAAACTTCAACCAATGTATTACTGAAAATAACAAATGTTGTAATTTCCCTGAGTGT 720
Db      590 TTTTAAACTTCAACCAATGTATTACTGAAAATAACAAATGTTGTAATTTCCCTGAGTGT 649
Qy      721 TATCTACTTGTATTAAGGTAAATATACATATCAATTAATAATCTCAGGAGTCAATTGCC 780
Db      650 TATCTACTTGTATTAAGGTAAATATACATATCAATTAATAATCTCAGGAGTCAATTGCC 709
Qy      781 AGAGATTTTGGGGAGGAAATGTTATCAACGGTTTCATTGAAATTAATATCCAAAAGTT 840
Db      710 AGAGATTTTGGGGAGGAAATGTTATCAACGGTTTCATTGAAATTAATATCCAAAAGTT 769
Qy      841 ATTTCTCAGAAAAATCAATATAAGTTTGCATGTTTTTATTCTTAAACAATTTTAAAAA 900
Db      770 ATTTCTCAGAAAAATCAATATAAGTTTGCATGTTTTTATTCTTAAACAATTTTAAAAA 829
Qy      901 CCACTGTAGATGATGTAATAGGAGCTGTGACGATTTCTGCATATATCTATATAAATTA 960
Db      830 CCACTGTAGATGATGTAATAGGAGCTGTGACGATTTCTGCATATATCTATATAAATTA 889
Qy      961 TTAATAAGTCAATCAGTATTCAACATCTTTTACACTAAAAAGCC 1004
Db      890 TTAATAAGTCAATCAGTATTCAACATCTTTTACACTAAAAAGCC 933

RESULT 11
AF527553
LOCUS      Homo sapiens B melanoma antigen variant e (BAGE1) mRNA, complete
DEFINITION      cds, alternatively spliced.
ACCESSION      AF527553
VERSION      AF527553.1 GI:28173035
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1068)
AUTHORS      Ruault,M., Van Der Bruggen,P., Brun,M.E., Boyle,S., Roizes,G. and

```



Sario,A.D.  
New BAGE (B melanoma antigen) genes mapping to the juxtaacentromeric  
regions of human chromosomes 13 and 21 have a cancer/testis  
expression profile  
Eur. J. Hum. Genet. 10 (12), 833-840 (2002)  
22349465  
12461691  
2 (bases 1 to 1068)  
Roizès,G. and De Sario,A.  
Roizès,G. and De Sario,A.  
BAGE, a gene family generated by juxtaacentromeric reshuffling in  
the Homiidae lineage, is under selective pressure  
Genomics (2003) In press  
3 (bases 1 to 1068)  
Roizès,M., van der Bruggen,P., Brun,M.-E., Boyle,S., Roizès,G. and  
De Sario,A.  
Direct Submission  
Submitted (08-JUL-2002) Institut de Genetique Humaine, CNRS UPR  
1142, 141, rue de la Cardonille, Montpellier 34396, France  
Location/Qualifiers  
1..1068  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="13"  
1..1068  
/gene="BAGE1"  
201..332  
/gene="BAGE1"  
/notes="tumor antigen; alternatively spliced"  
/codon\_start=1  
/product="B melanoma antigen variant e"  
/protein\_id="AA032673.1"  
/db\_xref="GI:128173036"  
/translations="MAARAVFLSAQLLQARLMKEESPVSWRLEPEDGTALCFIF"  
ORIGIN  
Query Match 56.1%; Score 563.4; DB 9; Length 1068;  
Best Local Similarity 90.5%; Pred. No. 1.5e-120;  
Matches 64; Conservative 0; Mismatches 11; Indels 57; Gaps 2;  
QY 1 CGCAATTAGGTTCTCCGGTATCTCCCGTGTAGCTGCTCTGTTCCTCCGGCTTAGAGGACC 60  
DB 1 CGCAATTAGGTTCTCCGGTATCTCCCGTGTAGCTGCTCTGTTCCTCCGGCTTAGAGGACC 60  
QY 61 AGGAGAGGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGTCTGTCTCACTCTGGATG 120  
DB 61 AGGAGAGGGGGAGCTGGAGCTGGAGCTGTAAACCGTGGTCTGTCTCACTCTGGATG 120  
QY 121 GTGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTAGGAGGGCGGCTGAGCGGTAGG 180  
DB 121 GTGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTAGGAGGGCGGCTGAGCGGTAGG 180  
QY 181 AGTGGGCTGGAGCAGTAAATGGCGCCAGAGCGGTTTCTGGCATTGTCTGCCAGC 240  
DB 181 AGTGGGCTGGAGCAGTAAATGGCGCCAGAGCGGTTTCTGGCATTGTCTGCCAGC 240  
QY 241 TGCTCCAAGCCAGCTGATCAAGGAGGAGTCCCTGTGGTGGTGGAGTGGAGCTGGAGCTG 300  
DB 241 TGCTCCAAGCCAGCTGATCAAGGAGGAGTCCCTGTGGTGGTGGAGTGGAGCTGGAGCTG 300  
QY 301 AAGACGGCACAGCTCTGTGCTTCTATCTCTGAGGTTGTGGCAGCCAGCTGTAGGAGAGC 360  
DB 301 AAGACGGCACAGCTCTGTGCTTCTATCTCTGAGGTTGTGGCAGCCAGCTGTAGGAGAGC 360  
QY 361 GCAGCTCAACAGGAGCAATAGGAG-----GAGATGGAGTTTCACT 400  
DB 361 GCAGCTCAACAGGAGCAATAGGAGGGTAAACCGTGGAGGCCAAGTGGAGTGGAGTTTCACT 420  
QY 401 GTGTACCCAGGATGGTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 460  
DB 421 GTGTACCCAGGATGGTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTG 480

QY 461 AGTGCCGAGATTA-----CAGCGATGTG 483  
DB 481 AGTGCCGAGATTAAGGTTTATGAAGAGATTTCTCTCTTACCCCTTTAGCGATGTG 540  
QY 484 CATTTTGTAAGCACTTTGAGCCCACTATCAAAATGCTGTGAAGAGAAATGTACCAGATGT 543  
DB 541 CATTTTGTAAGCACTTTGAGCCCACTATCAAAATGCTGTGAAGAGAAATGTACCAGATGT 600  
QY 544 ATCATATATCTTGTGCTGAGGAGCGGCTCTCTTTCAGGATTTTCACTCAGATCTTCTGTC 603  
DB 601 ATCATATATCTTGTGCTGAGGAGCGGCTCTCTTTCAGGATTTTCACTCAGATCTTCTGTC 660  
QY 604 TTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTACGATAG 658  
DB 661 TTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACTTCTCTCTTCTG 715  
RESULT 12  
AC064811/c  
LOCUS  
DEFINITION  
Homo sapiens chromosome 4 clone RP11-638N24 map 4, WORKING DRAFT  
SEQUENCE, 38 unordered pieces.  
AC064811  
AC064811.2 GI:8018040  
VERSION  
HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 177896)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 4, clone RP11-638N24  
Unpublished  
2 (bases 1 to 177896)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schautet,S., Severy,P., Spencer,B.,  
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (22-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 22, 2000 this sequence version replaced gi:7637317.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIER  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9935  
Center clone name: 638 N.24  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731  
 Consensus quality: 157343 bases at least Q40  
 Consensus quality: 16922 bases at least Q30  
 Consensus quality: 171194 bases at least Q20  
 Insert size: 185000; agarose-fp  
 Insert size: 174136; sum-of-contigs  
 Quality coverage: 3.5 in Q20 bases; agarose-fp  
 Quality coverage: 3.7 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 38 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

* 1
* 1313: contig of 1313 bp in length
* 1314
* 1413: gap of 100 bp
* 1414
* 2682: contig of 1269 bp in length
* 2683
* 2782: gap of 100 bp
* 2783
* 4276: contig of 1494 bp in length
* 4277
* 4376: gap of 100 bp
* 4377
* 5635: contig of 1259 bp in length
* 5636
* 5735: gap of 100 bp
* 5736
* 7093: contig of 1358 bp in length
* 7094
* 7193: gap of 100 bp
* 7194
* 9078: contig of 1885 bp in length
* 9079
* 9178: gap of 100 bp
* 9179
* 10411: contig of 1237 bp in length
* 10416
* 10515: gap of 100 bp
* 12355: contig of 1840 bp in length
* 12356
* 12455: gap of 100 bp
* 12456
* 15119: contig of 2664 bp in length
* 15120
* 15129: gap of 100 bp
* 18382: contig of 3163 bp in length
* 18383
* 18482: gap of 100 bp
* 18483
* 19774: contig of 1292 bp in length
* 19775
* 19874: gap of 100 bp
* 20901: contig of 1027 bp in length
* 20902
* 21001: gap of 100 bp
* 21002
* 22663: contig of 1662 bp in length
* 22664
* 22763: gap of 100 bp
* 22764
* 24922: contig of 2164 bp in length
* 24928
* 25027: gap of 100 bp
* 25028
* 27690: contig of 2663 bp in length
* 27691
* 27790: gap of 100 bp
* 27791
* 30074: contig of 2284 bp in length
* 30075
* 30174: gap of 100 bp
* 30175
* 33195: contig of 3021 bp in length
* 33196
* 33295: gap of 100 bp
* 33296
* 34897: contig of 1602 bp in length
* 34898
* 34997: gap of 100 bp
* 34998
* 38603: contig of 3606 bp in length
* 38604
* 38703: gap of 100 bp
* 38704
* 42437: contig of 3734 bp in length
* 42438
* 42537: gap of 100 bp
* 42538
* 45993: contig of 4456 bp in length
* 45994
* 47093: gap of 100 bp
* 47094
* 51619: contig of 4526 bp in length
* 51719: gap of 100 bp
* 51720
* 55359: contig of 3640 bp in length
* 55360
* 55459: gap of 100 bp
* 55460
* 60487: contig of 5028 bp in length
* 60488
* 60587: gap of 100 bp
* 60588
* 64603: contig of 4022 bp in length
* 64610
* 64709: gap of 100 bp
* 64710
* 69559: contig of 4850 bp in length
* 69560
* 69659: gap of 100 bp
* 69660
* 74140: contig of 4481 bp in length
* 74141
* 74240: gap of 100 bp
* 74241
* 78594: contig of 4354 bp in length
* 78595
* 78694: gap of 100 bp

```

```

* 78695
* 83792: contig of 5098 bp in length
* 83793
* 83892: gap of 100 bp
* 83893
* 89496: contig of 5604 bp in length
* 89497
* 89596: gap of 100 bp
* 89597
* 92998: contig of 7702 bp in length
* 92999
* 97398: gap of 100 bp
* 97399
* 105662: contig of 8264 bp in length
* 105663
* 105762: gap of 100 bp
* 105763
* 114041: contig of 8279 bp in length
* 114042
* 114141: gap of 100 bp
* 114142
* 122099: contig of 7938 bp in length
* 122100
* 122199: gap of 100 bp
* 130991: contig of 8792 bp in length
* 130992
* 131091: gap of 100 bp
* 131092
* 143872: contig of 12781 bp in length
* 143873
* 143972: gap of 100 bp
* 143973
* 158554: contig of 11882 bp in length
* 158555
* 159544: gap of 100 bp
* 159555
* 177896: contig of 21942 bp in length.
*
* FEATURES
* source
* 1. 177896
* Location/Qualifiers
*   /organism="Homo sapiens"
*   /mol_type="genomic DNA"
*   /db_xref="taxon:9606"
*   /map="4"
*   /chromosome="4"
*   /clone="RP11-638N24"
*   /clone_lib="RPC1-11 Human Male BAC"
* 1. 1313
*   /note="assembly_fragment"
* 1414. 2682
*   /note="assembly_fragment"
* 2783. 4276
*   /note="assembly_fragment"
* 4377. 5635
*   /note="assembly_fragment"
* 5736. 7093
*   /note="assembly_fragment"
* 7194. 9078
*   /note="assembly_fragment"
* 9179. 10415
*   /note="assembly_fragment"
* 10516. 12355
*   /note="assembly_fragment"
* 12456. 15119
*   /note="assembly_fragment"
* 15220. 18382
*   /note="assembly_fragment"
* 18483. 19774
*   /note="assembly_fragment"
*   clone_end.T7
*   vector_side:right"
* 19875. 20901
*   /note="assembly_fragment"
* 21002. 22663
*   /note="assembly_fragment"
* 22764. 24927
*   /note="assembly_fragment"
* 25028. 27690
*   /note="assembly_fragment"
* 27791. 30074
*   /note="assembly_fragment"
* 30175. 33195
*   /note="assembly_fragment"
* 33296. 34897
*   /note="assembly_fragment"
*   clone_end.SP6
*   vector_side:left"
* 34998. 38603
*   /note="assembly_fragment"
* 38704. 42437
*   /note="assembly_fragment"
* 42538. 45993

```



Query Match		53.0%; Score 532.4; DB 2; Length 177896;
Best Local Similarity		99.8%; Pred. No. 3.2e-113;
Matches 533; Conservative		0; Mismatches 1; Indels 0; Gaps 0;
Qy	471	TTACAGGATGTGCAATTTGTAGACATTTGGAGCCACATATCAAAATGCTGTGAAGAGAAA 530
Db	132184	TTTACAGGATGTGCAATTTGTAGACATTTGGAGCCACATATCAAAATGCTGTGAAGAGAAA 132125
Qy	531	TGTACCCAGATGTATCATTTCTTGTGCTGCGAGGAGCGGCTCTTTCAGGATTTTCAGT 590
Db	132124	TGTACCCAGATGTATCATTTCTTGTGCTGCGAGGAGCGGCTCTTTCAGGATTTTCAGT 132065
Qy	591	CACATCTTCCTGCTTTGTCAGAACACATTTGACCAAGCTCTTCAAGAGATGTAAGTTTACT 650
Db	132064	CACATCTTCCTGCTTTGTCAGAACACATTTGACCAAGCTCTTCAAGAGATGTAAGTTTACT 132005
Qy	651	ACGCATAGACTTTTAACTTCAACCATGTATTTACTGAAATTAACAATTTGTAAT 710
Db	132004	ACGCATAGACTTTTAACTTCAACCATGTATTTACTGAAATTAACAATTTGTAAT 131945
Qy	711	CCCTGAGTGTATTTCTACTTGTATTTAAAGGTAAATATACATATCAATTAATAATCTGAGG 770
Db	131944	CCCTGAGTGTATTTCTACTTGTATTTAAAGGTAAATATACATATCAATTAATAATCTGAGG 131885
Qy	771	GATCATTCAGAGATTTGTTGGGAGGAAATGTTATCAACGGTTTCATTTGAAATTAAT 830
Db	131884	GATCATTCAGAGATTTGTTGGGAGGAAATGTTATCAACGGTTTCATTTGAAATTAAT 131825
Qy	831	CCAAAAAGTTATTTCTCCAGAAAAATCAAAATAAGTTTGCATCTTTTATTTCTTAAAC 890
Db	131824	CCAAAAAGTTATTTCTCCAGAAAAATCAAAATAAGTTTGCATCTTTTATTTCTTAAAC 131765
Qy	891	ATTTTAAAAACCGTGTAGAAATGATTAATAGGAGCTGTGCAATTTCTGACATATAC 950
Db	131764	ATTTTAAAAACCGTGTAGAAATGATTAATAGGAGCTGTGCAATTTCTGACATATAC 131705
Qy	951	TATAAATTTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004
Db	131704	TATAAATTTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 131651
RESULT 13		
AC133563		
LOCUS		139063 bp DNA linear HTG 14-SEP-2002
DEFINITION		Homo sapiens chromosome 16 clone RP11-6218, WORKING DRAFT SEQUENCE, 15 unordered pieces.
ACCESSION		AC133563
VERSION		AC133563.1 GI:22857559
KEYWORDS		HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
SOURCE		Homo sapiens (human)
ORGANISM		Homo sapiens
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS		DOE Joint Genome Institute.
TITLE		Sequencing of Human Chromosome 16
JOURNAL		Unpublished
AUTHORS		DOE Joint Genome Institute.
TITLE		Direct Submission
JOURNAL		Submitted (14-SEP-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT		-----Genome Center Center: Joint Genome Institute Center Code: JGI Web site: <a href="http://www.jgi.doe.gov">http://www.jgi.doe.gov</a> ----- Project Information Center Project Name: 430373 Center clone name: RPCI-11_6218 ----- Summary Statistics

Consensus quality: 132123 bases at least Q40		
Consensus quality: 133813 bases at least Q30		
Consensus quality: 134589 bases at least Q20		
Estimated insert size: 172000; agarose-fp estimation		
Estimated insert size: 137663; sum-of-contigs estimation		
Quality coverage: 8.97 in Q20 bases; agarose-fp estimation		
Quality coverage: 11.21 in Q20 bases; sum-of-contigs estimation.		
* NOTE: This is a 'working draft' sequence. It currently		
* consists of 15 contigs. The true order of the pieces		
* is not known and their order in this sequence record is		
* arbitrary. Gaps between the contigs are represented as		
* runs of N, but the exact sizes of the gaps are unknown.		
* This record will be updated with the finished sequence		
* as soon as it is available and the accession number will		
* be preserved.		
* 1		
1475: contig of 1475 bp in length		
1476 1575: gap of unknown length		
1576 2662: contig of 1087 bp in length		
2663 2762: gap of unknown length		
2763 4032: contig of 1270 bp in length		
4033 4132: gap of unknown length		
4133 5295: contig of 1163 bp in length		
5296 5395: gap of unknown length		
5396 7456: contig of 2061 bp in length		
7457 7556: gap of unknown length		
7557 10057: contig of 2501 bp in length		
10058 10157: gap of unknown length		
10158 16520: contig of 6363 bp in length		
16521 16620: gap of unknown length		
16621 26759: contig of 10039 bp in length		
26760 26759: gap of unknown length		
26760 35377: contig of 8618 bp in length		
35378 35477: gap of unknown length		
35478 46059: contig of 10582 bp in length		
46060 46159: gap of unknown length		
46160 55012: contig of 8853 bp in length		
55013 55112: gap of unknown length		
55113 66795: contig of 11683 bp in length		
66796 66895: gap of unknown length		
66896 79813: contig of 12918 bp in length		
79814 79913: gap of unknown length		
79914 97772: contig of 17859 bp in length		
97773 97872: gap of unknown length		
97873 139063: contig of 41191 bp in length.		
Location/Qualifiers		
1..139063		
/organism="Homo sapiens"		
/mol_type="genomic DNA"		
/db_xref="taxon:9606"		
/chromosome="16"		
/clone="RP11-6218"		
/clone_lib="RPCI human BAC library 11"		
FEATURES		
source		
Query Match		
Best Local Similarity		
Matches 531; Conservative		
0; Mismatches 3; Indels 0; Gaps 0;		
Qy	471	TTACAGGATGTGCAATTTGTAGACATTTGGAGCCACATATCAAAATGCTGTGAAGAGAAA 530
Db	86678	TTTACAGGATGTGCAATTTGTAGACATTTGGAGCCACATATCAAAATGCTGTGAAGAGAAA 86737
Qy	531	TGTACCCAGATGTATCATTTCTTGTGCTGCGAGGAGCGGCTCTTTCAGGATTTTCAGT 590
Db	86738	TGTACCCAGATGTATCATTTCTTGTGCTGCGAGGAGCGGCTCTTTCAGGATTTTCAGT 86797
Qy	591	CACATCTTCCTGCTTTGTCAGAACACATTTGACCAAGCTCTTCAAGAGATGTAAGTTTACT 650
Db	86798	CACATCTTCCTGCTTTGTCAGAACACATTTGACCAAGCTCTTCAAGAGATGTAAGTTTACT 86857
Qy	651	ACGCATAGACTTTTAACTTCAACCATGTATTTACTGAAATTAACAATTTGTAAT 710
Db	86858	ACGCATAGACTTTTAACTTCAACCATGTATTTACTGAAATTAACAATTTGTAAT 86917

```

Qy 711 CCTGAGTGTATTCTACTTGTATTAAAGGTAATAATACATATCAATTAATAATCTGAGG 770
Db 86918 CCTGAGTGTATTCTACTTGTATTAAAGGTAATAATACATATCAATTAATAATCTGAGG 86977
Qy 771 GATCATGCCAGAGATTGTTGGGAGGGAATGTTATCAACGTTTCATTGAATTAAT 830
Db 86978 GATCGTTCAGAGATTGTTGGGAGGGAATGTTATCAACGTTTCATTGAATTAAT 87037
Qy 831 CCAAAAAGTTATTTCCCTCAGAAAAATCAATAAAGTTTGCAATTTTATTTTAAAC 890
Db 87038 CCAAAAAGTTATTTCCCTCAGAAAAATCAATAAAGTTTGCAATTTTATTTTAAAC 87097
Qy 891 ATTTTAAACCACTGTAGATGATGTAAATAGGCTGTCAGTATTTCTGCATATAC 950
Db 87098 ATTTTAAACCACTGTAGATGATGTAAATAGGCTGTCAGTATTTCTGCATATAC 87157
Qy 951 TATAAATATTATAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004
Db 87158 TATAAATATTATAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 87211

RESULT 14
AC145613 178300 bp DNA linear HTG 19-JUL-2003
LOCUS Homo sapiens chromosome UNK clone RP11-1266H24, *** SEQUENCING IN
DEFINITION PROGRESS ***, 14 unordered pieces.
ACCESSION AC145613
VERSION AC145613.1 GI:32996952
KEYWORDS HTG; HTGS PHASE1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178300)
Wilson, R.K.
The sequence of Homo sapiens clone
Unpublished
REFERENCE 2 (bases 1 to 178300)
Wilson, R.K.
Direct Submission
Submitted (19-JUL-2003) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH1266H24
----- Summary Statistics -----
Sequencing vector: p13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170279 bases at least Q40
Consensus quality: 171980 bases at least Q30
Consensus quality: 173066 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1534: contig of 1534 bp in length
* 1535 1634: gap of unknown length
* 1635 4097: contig of 2463 bp in length
* 4098 4197: gap of unknown length

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH1266H24
----- Summary Statistics -----
Sequencing vector: p13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170279 bases at least Q40
Consensus quality: 171980 bases at least Q30
Consensus quality: 173066 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1534: contig of 1534 bp in length
* 1535 1634: gap of unknown length
* 1635 4097: contig of 2463 bp in length
* 4098 4197: gap of unknown length

FEATURES
source
1..178300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UNK"
/clones="RP11-1266H24"
1..1534
/note="assembly_name:Contig13"
1535..4097
/note="assembly_name:Contig14"
4098..6161
/note="assembly_name:Contig15"
6162..11394
/note="assembly_name:Contig16"
11395..13311
/note="assembly_name:Contig17"
13312..16041
/note="assembly_name:Contig18"
16042..26488
/note="assembly_name:Contig19"
26489..47629
/note="assembly_name:Contig20"
47630..70793
/note="assembly_name:Contig21"
70794..102752
/note="assembly_name:Contig22"
102753..136872
/note="assembly_name:Contig23"
136873..174989
/note="assembly_name:Contig24"
174990..176837
/note="assembly_name:Contig25"
176838..178300
/note="assembly_name:Contig26"
178301..178300
/note="assembly_name:Contig27"

Location/Qualifiers
1..178300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UNK"
/clones="RP11-1266H24"
1..1534
/note="assembly_name:Contig13"
1535..4097
/note="assembly_name:Contig14"
4098..6161
/note="assembly_name:Contig15"
6162..11394
/note="assembly_name:Contig16"
11395..13311
/note="assembly_name:Contig17"
13312..16041
/note="assembly_name:Contig18"
16042..26488
/note="assembly_name:Contig19"
26489..47629
/note="assembly_name:Contig20"
47630..70793
/note="assembly_name:Contig21"
70794..102752
/note="assembly_name:Contig22"
102753..136872
/note="assembly_name:Contig23"
136873..174989
/note="assembly_name:Contig24"
174990..176837
/note="assembly_name:Contig25"
176838..178300
/note="assembly_name:Contig26"
178301..178300
/note="assembly_name:Contig27"

Query Match 52.4%; Score 526; DB 2; Length 178300;
Best Local Similarity 99.1%; Pred. No. 9.8e-112;
Matches 529; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 471 TTACGCGATGTCATTTTGAAGCACTTTGAGCCACTATCAATGCTGTGAAGAGAA 530
Db 165850 TTACGCGATGTCATTTTGAAGCACTTTGAGCCACTATCAATGCTGTGAAGAGAA 165909
Qy 531 TGTACCCAGATGTCATTTATTCCTTGTGCTCAGAGCCGCTCTCTTTCAGGATTTCACT 590
Db 165910 TGTACCCAGATGTCATTTATTCCTTGTGCTCAGAGCCGCTCTCTTTCAGGATTTCACT 165969
Qy 591 CACATCTTCTGCTTTGTCGAGAACACATTCACCAAGCTCTCTGAAAGATGTAAGTTTACT 650

```

```
Db 165970 CACATCTTCCTGCTTGTCCAGAACACATGACCAAGCTCTGAAAGATGTAAGTTTACT 166029
Qy 651 ACGCATAGACTTTTAAACTTCAACCAATGATTTTACTGAAATAAACAATGTTGTAAT 710
Db 166030 ACGCATAGACTTTTAAACTTCAACCAATGATTTTACTGAAATAAACAATGTTGTAAT 166089
Qy 711 CCCTGAGTGTTATCTACTTGTATTAAGAGTAATATACATATCAATCAATTAATCTGAGG 770
Db 166090 CCCTGAGTGTTATCTACTTGTATTAAGAGTAATATACATATCAATCAATTAATCTGAGG 166149
Qy 771 GATCATTGCCAGAGATTTGGGGAGGAAATGTTATCAACCGTTCATTGAAATTAAT 830
Db 166150 GATCGTTGCCAGAGATTTGGGGAGGAAATGTTATCAACCGTTCATTGAAATTAAT 166209
Qy 831 CCAAAAAGTTATTTCCCTCAGAAAATCAATAAAGTTTGCATGTTTTTATCTTAAAC 890
Db 166210 CCAAAAAGTTATTTCCCTCAGAAAATCAATAAAGTTTGCATGTTTTTATCTTAAAC 166269
Qy 891 ATTTTAAAAACCACTGTAGATGATGTAATAGGACTGTGCAATTTCTGACATATAC 950
Db 166270 ATTTTAAAAACCACTGTAGAAAGATGTAATAGGACTGTGCAATTTCTGACATATAC 166329
Qy 951 TATTAATTTTAAAAAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004
Db 166330 TATTAATTTTAAAAAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 166383

RESULT 15
AF155875/c
LOCUS Homo sapiens clone MC169, *** SEQUENCING IN PROGRESS ***, 2 ordered
DEFINITION pieces.
ACCESSION AF155875
VERSION AF155875.1 GI:18000259
KEYWORDS HTG; HTGS PHASE2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107701)
Xia, J.-H., Huang, L., Tang, D.-S., Dai, H.-P., Pan, Q. and Long, Z.-G.
Direct Submission
Submitted (18-MAY-1999) National Lab of Medical Genetics of China,
Hunan Medical University, 88 Xiang-ya Road, Changsha, Hunan 410078,
P.R. China
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 105436: contig of 105436 bp in length
* 105437 105536: gap of unknown length
* 105537 107701: contig of 2165 bp in length.
FEATURES
Location/Qualifiers
1..107701
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"
/clone="MC169"
ORIGIN
Query Match 52.2%; Score 524.4; DB 2; Length 107701;
Best Local Similarity 98.9%; Pred. No. 2.3e-111;
Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 471 TTACAGCGATGTCATTTGTAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAA 530
Db 77748 TTTACGAATGTGCATTTTGTAAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAA 77689
```

Search completed: July 9, 2004, 08:57:30  
Job time : 6481 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 23:14:12 ; Search time 633 Seconds  
(without alignments)  
6738.058 Million cell updates/sec

Title: US-10-081-108-1

Perfect score: 1004

Sequence: 1 CGCCAAATTAGGCTCTCCG.....ATCTTTACACTAAAGGCC 1004

Scoring table: IDENTITY NJC

Gapop 10.0 , Gapext 1.0

Searched: 3379863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004	100.0	1004	2	AAT36382
2	1004	100.0	1004	2	ABQ76214
3	1004	100.0	1004	6	ABV72285
4	1004	100.0	1004	9	ADD25519
5	958	95.4	1032	2	AAQ81005
6	524.4	52.2	11162	6	ABL55839
7	524.4	52.2	11162	6	ABL54504
8	524.4	52.2	11162	6	AAK99658
9	524.4	52.2	11162	6	ABA99937
10	524.4	52.2	13928	6	AAH77500
11	524.4	52.2	107602	6	AAK99657
12	524.4	52.2	107612	6	ABL54503
13	181.8	18.1	57728	4	AAK87588
14	149.8	14.9	6292	4	AAK46735
15	106.2	10.6	251	3	AAK19397
16	87	8.7	130	4	AAK136471
17	87	8.7	130	7	ABX59459
18	87	8.7	2028	5	ABV22788
19	87	8.7	2028	5	ABV28615
20	87	8.7	21423	4	AAK136470
21	87	8.7	21423	7	ABX59458
22	86.6	8.6	476	6	ABV87857
23	86.2	8.6	72332	8	ADA02552

C	24	86.2	8.6	72332	9	ADB72290	ADB72290 Human WNT
C	25	86	8.6	6292	4	AS45736	AS45736 Tumour su
C	26	85.6	8.5	411	6	ABL66140	ABL66140 Lung can
C	27	85.6	8.5	411	6	ABN93947	ABN93947 Gene #445
C	28	85.6	8.5	113000	8	ABT44365	ABT44365 Partial g
C	29	85.6	8.5	325791	4	AAK43104	AAK43104 Human Oes
C	30	85.4	8.5	4048	7	ABZ68064	ABZ68064 Human sec
C	31	85.4	8.5	4048	7	ABZ74530	ABZ74530 Secreted
C	32	85.4	8.5	4048	9	ADC20955	ADC20955 Human sec
C	33	85.2	8.5	39287	6	ABN80533	ABN80533 Human P45
C	34	84.4	8.4	378	5	ABV15229	ABV15229 Human pro
C	35	84.4	8.4	501	5	ABV49004	ABV49004 Human pro
C	36	84.4	8.4	667	4	AAK03915	AAK03915 Human rep
C	37	84.4	8.4	667	4	AAK03913	AAK03913 Human rep
C	38	84.4	8.4	667	5	AAK40313	AAK40313 DNA encod
C	39	84.4	8.4	667	5	AAK40315	AAK40315 DNA encod
C	40	84.2	8.4	249	3	AAK12905	AAK12905 Human sec
C	41	84	8.4	300	4	AAK68680	AAK68680 Human imm
C	42	84	8.4	358	8	ACH42606	ACH42606 Human pro
C	43	84	8.4	456	5	ABV59611	ABV59611 Human foe
C	44	84	8.4	571	6	ABN61772	ABN61772 Human can
C	45	84	8.4	690	4	AAK78891	AAK78891 Human imm

## ALIGNMENTS

## RESULT 1

AAT36382

ID AAT36382 standard; cDNA; 1004 BP.

XX AC AAT36382;

DT 04-DEC-1996 (first entry)

DE BAGE tumour rejection antigen precursor cDNA.

KW BAGE; tumour rejection antigen precursor; TRAP; MHC;

KW major histocompatibility complex; HLA-Cw\*1601; melanoma; metastasis;

KW diagnosis; therapy; vaccine; ss.

XX Homo sapiens.

OS Homo sapiens.

FH Key Location/Qualifiers

FT primer\_bind Complement(100..117)

FT /tag= a

FT /note= "primer VDB85"

FT 201..332

FT /tag= b

FT /product= "tumour rejection antigen precursor"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

FT 367..385

FT /tag= d

FT /note= "primer VDB86"

FT 385..484

FT /tag= e

FT /note= "Alu repeat"

FT 204..230

FT /tag= c

FT /product= "tumour rejection antigen"

DR WPI; 1996-393411/39.  
 XX P-PSDB; AAW02152.  
 PT Tumour rejection antigen precursor (TRAP) and gene - useful to develop  
 PT prods. for diagnosis and treatment of disorders characterised by TRAP,  
 PT partic. melanomas.  
 XX  
 PS Claim 1; Page 27-28; 44pp; English.  
 XX  
 CC A cDNA clone (AAT36382), designated cDNA-AD5, codes for the BAGE tumour  
 CC rejection antigen precursor (TRAP) (AAW02152). MHC molecule HLA-Cw\*1601  
 CC presents a tumour rejection antigen derived from the BAGE TRAP. The cDNA  
 CC was identified by cotransfecting HLA-Cw\*1601 cDNA with a cDNA library  
 CC derived from melanoma cell line M22-MEL.43 into COS-7 cells and isolating  
 CC clones capable of stimulating prodn. of tumour necrosis factor by  
 CC cytotoxic T-lymphocytes. In 600 samples of tumours, the BAGE gene was  
 CC expressed mainly in melanomas (esp. in metastatic lesions), bladder  
 CC carcinomas and mammary carcinomas. The isolated gene may be used  
 CC diagnostically or for prodn. of the BAGE TRAP  
 XX  
 SQ Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;

Query Match 100.0%; Score 1004; DB 2; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-250;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGCCAAATTTAGGCTCTCCGGTATCTCCCGTAGCTGCTGTTCCTCCGGCTTAGAGACC 60  
 DB 1 OGCCAAATTTAGGCTCTCCGGTATCTCCCGTAGCTGCTGTTCCTCCGGCTTAGAGACC 60

QY 61 AGGAGAGGGGGAGCTGGAGGCTGGAGGCTGTGAACACCGTGGCTGTCTCACTTGGATG 120  
 DB 61 AGGAGAGGGGGAGCTGGAGGCTGGAGGCTGTGAACACCGTGGCTGTCTCACTTGGATG 120

QY 121 GTGGTGCACACAGATGGCAGCGCAGCTGGAGTGTAGAGGGCGGCGCTGAGCGGTAGG 180  
 DB 121 GTGGTGCACACAGATGGCAGCGCAGCTGGAGTGTAGAGGGCGGCGCTGAGCGGTAGG 180

QY 181 AGTGGGGCTGGAGCAGTAAGATGGCGGCGCAGAGCGGTTTTCTGGCATTTGTCTGCCACG 240  
 DB 181 AGTGGGGCTGGAGCAGTAAGATGGCGGCGCAGAGCGGTTTTCTGGCATTTGTCTGCCACG 240

QY 241 TGCTTCAAGCCAGCTGATGAAGAGAGAGTCCCTGTGTGAGCTGGAGGTTGGAGCCGTG 300  
 DB 241 TGCTTCAAGCCAGCTGATGAAGAGAGAGTCCCTGTGTGAGCTGGAGGTTGGAGCCGTG 300

QY 301 AAGACGGCAGAGCTCTGTGCTTCTCATCTTGTAGGTTGTGGCAGCCAGCTGTAGGAGCG 360  
 DB 301 AAGACGGCAGAGCTCTGTGCTTCTCATCTTGTAGGTTGTGGCAGCCAGCTGTAGGAGCG 360

QY 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGAGTTCCTTCTGAGGTTGTGGCAGCCAGCTGTAGGAGCG 420  
 DB 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGAGTTCCTTCTGAGGTTGTGGCAGCCAGCTGTAGGAGCG 420

QY 421 GATCTCTGACCTCGTATCGCCCGCTTGGCTTCCAAAGTCCGAGATTACACGAT 480  
 DB 421 GATCTCTGACCTCGTATCGCCCGCTTGGCTTCCAAAGTCCGAGATTACACGAT 480

QY 481 GTGCAATTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAGAGAAATGTACCAGAG 540  
 DB 481 GTGCAATTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAGAGAAATGTACCAGAG 540

QY 541 TGATCATTTATCTTGTGCTGAGGAGCGGCTTTCAGGATTTCACTACATCTTCC 600  
 DB 541 TGATCATTTATCTTGTGCTGAGGAGCGGCTTTCAGGATTTCACTACATCTTCC 600

QY 601 TGCTTTTCCAGAACACATTTGACCAAGCTCTCTGAAGAGATGTAAGTTTACTACGATAGAC 660  
 DB 601 TGCTTTTCCAGAACACATTTGACCAAGCTCTCTGAAGAGATGTAAGTTTACTACGATAGAC 660

QY 661 TTTTAAACTTCAACCAATGATTTACTGAAATTAACCAATGTTGTAATTCCTCGATGT 720  
 DB 661 TTTTAAACTTCAACCAATGATTTACTGAAATTAACCAATGTTGTAATTCCTCGATGT 720

QY 721 TATTCTACTTGTATTAAAGGTAATTAATACATTAATCAATTAATCTGAGGATCATTTGCC 780  
 DB 721 TATTCTACTTGTATTAAAGGTAATTAATACATTAATCAATTAATCTGAGGATCATTTGCC 780

QY 781 AGAGATTGTTGGGAGGAGAAATGTTATCAACGGTTCATTGAAATTAATCCAAAAGTT 840  
 DB 781 AGAGATTGTTGGGAGGAGAAATGTTATCAACGGTTCATTGAAATTAATCCAAAAGTT 840

QY 841 ATTTCTCTCAGAAAATCAAAATAAAGTTTGCATGTTTTTAACTTTTAAACATTTTAAAA 900  
 DB 841 ATTTCTCTCAGAAAATCAAAATAAAGTTTGCATGTTTTTAACTTTTAAACATTTTAAAA 900

QY 901 CCACCTGTAGATGATGTAATAATAGGAGCTGTGCAGTATTTCTGACATATATAATAATTA 960  
 DB 901 CCACCTGTAGATGATGTAATAATAGGAGCTGTGCAGTATTTCTGACATATATAATAATTA 960

QY 961 TTAATAAGTCAATCAGTATTCAACATCTTTTACACTTAAAGGCC 1004  
 DB 961 TTAATAAGTCAATCAGTATTCAACATCTTTTACACTTAAAGGCC 1004

RESULT 2  
 ABO76214  
 ID ABO76214 standard; DNA; 1004 BP.  
 XX  
 AC ABO76214;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Human tumour antigen BAGE DNA.  
 XX  
 KW Tumour antigen; human; vaccine; cellular immune response; immunogen;  
 KW cancer; tumour; BAGE; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6287569-B1.  
 XX  
 PD 11-SEP-2001.  
 XX  
 PF 06-APR-1998; 98US-00056105.  
 XX  
 PR 10-APR-1997; 97US-0043467P.  
 XX  
 PA (REGC) UNIV CALIFORNIA.  
 XX  
 PI Kipps TJ, Wu Y;  
 XX  
 DR WPI; 1998-583198/49.  
 XX  
 CC Generating cellular immune response in patient to target protein -  
 CC comprises introducing vector with nucleotide sequence encoding immunogen  
 CC comprising protein processing signal into cell of patient.  
 PS Disclosure; Col 73-76; 61pp; English.  
 XX  
 CC This invention describes a novel method for generating a cellular immune  
 CC response in a patient to a target protein or its fragment. The method  
 CC involves introducing a vector containing a nucleotide sequence encoding a  
 CC chimeric immunogen comprising a protein processing signal and the target  
 CC protein or its fragment. The immunogen is produced by the cells and  
 CC processed so that the target protein or its fragment is presented to the  
 CC patients immune system and a cellular immune response is initiated. The  
 CC method and vectors can be used as a form of vaccination and could be used  
 CC to generate a cellular immune response in patients to, e.g. cancerous  
 CC tumours. The cellular immune response is the predominant immune response  
 CC in the patient. This sequence represents a DNA fragment which encodes the  
 CC human tumour antigen BAGE described in the method of the invention. Note:  
 CC The information in this spec has been previously disclosed in WO199845444  
 CC however this spec contained no sequence information  
 XX  
 SQ Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;

Query Match 100.0%; Score 1004; DB 2; Length 1004;  
Best Local Similarity 100.0%; Pred. No. 1.8e-250;  
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCAAATTTAGGGTCTCCGGTATCTCCCGTGTAGCTGCTGTTCCTCCGGCTTAGAGGACC 60  
DB 1 CGCCAAATTTAGGGTCTCCGGTATCTCCCGTGTAGCTGCTGTTCCTCCGGCTTAGAGGACC 60

QY 61 AGGAGAAAGGGGAGCTGGAGGCTGTGAACCCGTGTGAACCCGTGTGCTCTCACTCTGGATG 120  
DB 61 AGGAGAAAGGGGAGCTGGAGGCTGTGAACCCGTGTGAACCCGTGTGCTCTCACTCTGGATG 120

QY 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTAGGAGGCGGCTGACCGGTAGG 180  
DB 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTAGGAGGCGGCTGACCGGTAGG 180

QY 181 AGTGGGCTGGAGCAGTAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240  
DB 181 AGTGGGCTGGAGCAGTAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240

QY 241 TGCTCCAGCCAGCTGTGATGAAGAGAGTCCCTGTGTGAGCTGGAGTGGAGCCCTG 300  
DB 241 TGCTCCAGCCAGCTGTGATGAAGAGAGTCCCTGTGTGAGCTGGAGTGGAGCCCTG 300

QY 301 AAGAGGCGCAGCTGTGCTTCTCATCTCTGAGGTTGTGCGCAGCCACGGTGTAGGAGCG 360  
DB 301 AAGAGGCGCAGCTGTGCTTCTCATCTCTGAGGTTGTGCGCAGCCACGGTGTAGGAGCG 360

QY 361 GCAGCTCAACAGAGCAATAGGAGGATGGAGTTTCACTGTGTGAGCCAGGATGCTCTC 420  
DB 361 GCAGCTCAACAGAGCAATAGGAGGATGGAGTTTCACTGTGTGAGCCAGGATGCTCTC 420

QY 421 GATCTCTGACCTGTGATCCGCGCCGCTTGGCCCTTCCAAAGTCCGAGATTACAGCGAT 480  
DB 421 GATCTCTGACCTGTGATCCGCGCCGCTTGGCCCTTCCAAAGTCCGAGATTACAGCGAT 480

QY 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCAGA 540  
DB 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCAGA 540

QY 541 TGATCATATCTCTGTGTGAGGAGCGGCTCTTTCAGGATTTCACTCACTCTCC 600  
DB 541 TGATCATATCTCTGTGTGAGGAGCGGCTCTTTCAGGATTTCACTCACTCTCC 600

QY 601 TGCTTTGTCCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGCATAGAC 660  
DB 601 TGCTTTGTCCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGCATAGAC 660

QY 661 TTTTAACTTCAACCAATGTATTTACTGAAATAACAAATGTTGTAATTTCCCTGAGTGT 720  
DB 661 TTTTAACTTCAACCAATGTATTTACTGAAATAACAAATGTTGTAATTTCCCTGAGTGT 720

QY 721 TATTTCTATTTATTAAGGTAATATACATATCAATTAATCTGAGGGATCATGTC 780  
DB 721 TATTTCTATTTATTAAGGTAATATATCAATATCAATTAATCTGAGGGATCATGTC 780

QY 781 AGAGATTTTGGGGAGGAAATGTTATCAACGGTTTCATTGAAATTAATCCAAAAGTT 840  
DB 781 AGAGATTTTGGGGAGGAAATGTTATCAACGGTTTCATTGAAATTAATCCAAAAGTT 840

QY 841 ATTTCTCAGAAAATCAATAAAGTTTGCATGTTTTTATTTCTTAAACATTTTAAAAA 900  
DB 841 ATTTCTCAGAAAATCAATAAAGTTTGCATGTTTTTATTTCTTAAACATTTTAAAAA 900

QY 901 CCACTGTAGATGATGTAATAGGACTGTGCAATATTTCTGACATATATAATAATTA 960  
DB 901 CCACTGTAGATGATGTAATAGGACTGTGCAATATTTCTGACATATATAATAATTA 960

QY 961 TTTAAAGTCAATCAGTATTCACATCTTTTACACTTAAAGCC 1004  
DB 961 TTTAAAGTCAATCAGTATTTCAACATCTTTTACACTTAAAGCC 1004

## RESULT 3

ABV72285

ID ABV72285 standard; DNA; 1004 BP.

XX AC ABV72285;

XX DT 16-DEC-2002 (first entry)

XX Nucleotide sequence of human BAGE.

XX Human; IFIT-2; chronic myelogenous leukemia; LAGE-1; BAGE; DDB1; ETS2;  
KW PIASy; PIASx-alpha; PIASx-beta; DAPK3; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 201..332

XX FT /\*tag= a

XX FT /product= "BAGE"

XX PN WO200270747-A1.

XX PD 12-SEP-2002.

XX PP 01-MAR-2002; 2002WO-JP001901.

XX PR 01-MAR-2001; 2001JP-00056438.

XX PA (FUJI ) FUJISAWA PHARM CO LTD.

XX PI Mano H;

XX DR WPI; 2002-682911/73.

XX DR P-PSDB; ABE78347.

XX PT Measuring the expression profile of genes in a cell or tissue sample for  
diagnosis of chronic myelogenous leukemia and identification of agents  
for its treatment.

XX PS Example 3; Page 38-39; 97pp; Japanese.

XX CC The present sequence encodes human BAGE. The expression level of the gene  
is used in the method of the invention. The specification describes a  
method of examining chronic myelogenous leukemia. The method comprises  
measuring the expression level of a gene selected from IFIT-2, LAGE-1,  
BAGE, DDB1, ETS2, PIASy, PIASx-alpha, PIASx-beta and DAPK3, or  
determining the expression profile of a group of genes including one or  
more of these genes, in a cell or tissue sample from a chronic  
myelogenous leukemia patient. The method is used for the diagnosis,  
treatment and prevention of chronic myelogenous leukemia

SQ Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;

Query Match 100.0%; Score 1004; DB 6; Length 1004;

Best Local Similarity 100.0%; Pred. No. 1.8e-250;

Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCAAATTTAGGGTCTCCGGTATCTCCCGTGTAGCTGCTGTTCCTCCGGCTTAGAGGACC 60

DB 1 CGCCAAATTTAGGGTCTCCGGTATCTCCCGTGTAGCTGCTGTTCCTCCGGCTTAGAGGACC 60

QY 61 AGGAGAAAGGGGAGCTGGAGGCTGGAGCCTGTAAACCCGTGGCTCTCACTCTGGATG 120

DB 61 AGGAGAAAGGGGAGCTGGAGGCTGGAGCCTGTAAACCCGTGGCTCTCACTCTGGATG 120

QY 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTAGGAGGCGGCTTAGCGGTAGG 180

DB 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTAGGAGGCGGCTTAGCGGTAGG 180

QY 181 AGTGGGCTGGAGCAGTAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240

DB 181 AGTGGGCTGGAGCAGTAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240

QY 241 TGCTCCAGCCAGGCTGATGAAGAGAGTCCCTGTGCTGAGCTGGAGGTTGGAGCCTG 300  
 Db 241 TGCTCCAGCCAGGCTGATGAAGAGAGTCCCTGTGCTGAGCTGGAGGTTGGAGCCTG 300  
 QY 301 AAGACGGCAGCTGCTGCTTCATCTCTGAGGTTGTGGCAGCCACGGTGTAGAGAGC 360  
 Db 301 AAGACGGCAGCTGCTGCTTCATCTCTGAGGTTGTGGCAGCCACGGTGTAGAGAGC 360  
 QY 361 GCAGCTCAACAGAGCAATAGGAGGAGATGAGGTTTCACTGTGTACAGCAGATGCTTC 420  
 Db 361 GCAGCTCAACAGAGCAATAGGAGGAGATGAGGTTTCACTGTGTACAGCAGATGCTTC 420  
 QY 421 GATCTCTGACCTCGTGATCCGCCCGCTTGGCTTCCAAAGTGCAGGATTAACAGCGAT 480  
 Db 421 GATCTCTGACCTCGTGATCCGCCCGCTTGGCTTCCAAAGTGCAGGATTAACAGCGAT 480  
 QY 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
 Db 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
 QY 541 TGATCATATTATCTTGTGCTGCAGGAGCCGCTTCCAGGATTTTCACTCATCTTCC 600  
 Db 541 TGATCATATTATCTTGTGCTGCAGGAGCCGCTTCCAGGATTTTCACTCATCTTCC 600  
 QY 601 TGCTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAAGTTTACTACGCATAGAC 660  
 Db 601 TGCTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAAGTTTACTACGCATAGAC 660  
 QY 661 TTTTAACTTCAACCAATGATTTTACTGAAATAACAAATGTTGTAATTCCTGAGTGT 720  
 Db 661 TTTTAACTTCAACCAATGATTTTACTGAAATAACAAATGTTGTAATTCCTGAGTGT 720  
 QY 721 TATTTCTACTTGTATTAAGAGTAAATATACATATCAATTAATTAATTAATTAATTAAT 780  
 Db 721 TATTTCTACTTGTATTAAGAGTAAATATACATATCAATTAATTAATTAATTAATTAAT 780  
 QY 781 AGAGATTGTTGGGGGGAATGTTTCAACGGTTTCAATTAATTAATTAATTAATTAATTAAT 840  
 Db 781 AGAGATTGTTGGGGGGAATGTTTCAACGGTTTCAATTAATTAATTAATTAATTAATTAAT 840  
 QY 841 ATTTCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATCTTAAACATTTTAAAAA 900  
 Db 841 ATTTCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATCTTAAACATTTTAAAAA 900  
 QY 901 CCACGTAGATGATGTAATAGGACTGTGCAGTATTTCTGACATATCTATAAATTA 960  
 Db 901 CCACGTAGATGATGTAATAGGACTGTGCAGTATTTCTGACATATCTATAAATTA 960  
 QY 961 TTAAGAGTCAATCAGTATTAATCAACATCTTTTACACTAAAAAGCC 1004  
 Db 961 TTAAGAGTCAATCAGTATTAATCAACATCTTTTACACTAAAAAGCC 1004

RESULT 4

ADD25519 standard; DNA; 1004 BP.

ID ADD25519;

AC ADD25519;

DT 15-JAN-2004 (first entry)

DE Binding domain-immunoglobulin fusion protein-associated DNA #44.

ds: Binding domain; immunoglobulin; fusion protein; cytostatic;  
 antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
 neuroprotective; hinge region; immunoglobulin heavy chain;  
 CH2 constant region; CH3 constant region; IGH1; ADCC; complement fixation;  
 antibody dependent cell-mediated cytotoxicity; melanoma; sarcoma;  
 malignant condition; B-cell disorder; myasthenia gravis; Grave's disease;  
 rheumatoid arthritis; multiple sclerosis; autoimmune disease.

OS Unidentified.  
 XX US2003118592-A1.  
 XX 26-JUN-2003.  
 XX 25-JUL-2002; 2002US-00207655.  
 XX 17-JAN-2001; 2001US-0367358P.  
 XX 17-JAN-2002; 2002US-00053530.  
 XX 03-JUN-2002; 2002US-0385691P.  
 XX (GENE-) GENE-CRAFT INC.  
 XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
 XX WPI; 2003-801317/75.

XX New binding domain-immunoglobulin fusion protein, useful for treating a  
 PT subject having or suspected of having a malignant condition or a B-cell  
 PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.

XX Disclosure; SEQ ID NO 80; 157pp; English.

XX Unidentified

XX Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 1004; DB 9; Length 1004;

XX Best Local Similarity 100.0%; Pred. No. 1.8e-250;

XX Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCAAATTTAGGCTCGGTATCTCCCGTGTAGCTCTCTGTTCCTGGCTTAGAGGACC 60

Db 1 CGCCAAATTTAGGCTCGGTATCTCCCGTGTAGCTCTCTGTTCCTGGCTTAGAGGACC 60

QY 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACACCGTGGCTCTCACTCTGGATG 120

Db 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACACCGTGGCTCTCACTCTGGATG 120

QY 121 GTGTGGGCAACAGAGATGGAGCGGAGCTGTAGAGGCTGTAGAGGCGGCTGAGCGGTAGG 180

Db 121 GTGTGGGCAACAGAGATGGAGCGGAGCTGTAGAGGCTGTAGAGGCGGCTGAGCGGTAGG 180

QY 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTGTCTCCAGC 240

Db 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTGTCTCCAGC 240

QY 241 TGCTCCAAAGCCAGGCTGTATGAAGAGAGTCCCTGTGTGAGCTGGAGGTTGGAGCCTG 300

Db 241 TGCTCCAAAGCCAGGCTGTATGAAGAGAGTCCCTGTGTGAGCTGGAGGTTGGAGCCTG 300

QY 301 AAGACGGCAGCTCTGTGCTTCTATCTTCTGAGCTGTGGCAGCCACGGTGTAGAGAGC 360

Db 301 AAGACGGCAGCTCTGTGCTTCTATCTTCTGAGCTGTGGCAGCCACGGTGTAGAGAGC 360

QY 361 GCAGCTCAACAGGAGCAATAGGAGAGATGAGGTTTCACTGTGTAGCAGGAGTGGTCTC 420

Db 361 GCAGCTCAACAGGAGCAATAGGAGAGATGAGGTTTCACTGTGTAGCAGGAGTGGTCTC 420

QY 421 GATCTCTGACCTCGTGATCCGCCCGCTTGGCTTCCAAAGTGCAGGATTAACAGCGAT 480

Db 421 GATCTCTGACCTCGTGATCCGCCCGCTTGGCTTCCAAAGTGCAGGATTAACAGCGAT 480

QY 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540

Db 481 GTGCATTTTGAAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540

QY 541 TGTATCATTTTCTGCTGCTGCAGGAGCCGCTCTTTTCAAGATTTTCACTCATCTTCC 600

Db 541 TGTATCATTTTCTGCTGCTGCAGGAGCCGCTCTTTTCAAGATTTTCACTCATCTTCC 600

QY 601 TGCTTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAAGTTTACTACGCATAGAC 660

Db 601 TGCTTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAAGTTTACTACGCATAGAC 660



Db 601 TGCTTTGTCAGAACACATTGACCAAGCTCTGAAAGATGTAAGTTTACTACCATAGAC 660  
 QY 661 TTTTAAATTTCAACCAATGATTTACTGAAAAATAAACHAATGTTGTAAATTCCTGAGTGT 720  
 Db 661 TTTTAAATTTCAACCAATGATTTACTGAAAAATAAACHAATGTTGTAAATTCCTGAGTGT 720  
 QY 721 TATTCTACTTGTATTTAAAGTAATAATACATATCATTAATCAATTAATCTGAGGATCATTTGCC 780  
 Db 721 TATTCTACTTGTATTTAAAGTAATAATACATATCATTAATCAATTAATCTGAGGATCATTTGCC 780  
 QY 781 AGAGATTGTTGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAATPCCAAAAGTT 840  
 Db 781 AGAGATTGTTGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAATPCCAAAAGTT 840  
 QY 841 ATTTCCTCAGAAAAATCAAAATAAAGTTTGCATGTTTTTATTCTTAAACATTTTAAAAA 900  
 Db 841 ATTTCCTCAGAAAAATCAAAATAAAGTTTGCATGTTTTTATTCTTAAACATTTTAAAAA 900  
 QY 901 CCACCTGTAGAAATGATGTAATAGGACTGTGCAGTATTTCTGCATATATCTATATAATTA 960  
 Db 901 CCACCTGTAGAAATGATGTAATAGGACTGTGCAGTATTTCTGCATATATCTATATAATTA 960  
 QY 961 TTAATAAGTCAATCAGTATTCACATCTTTTACACTAAAAGCC 1004  
 Db 961 TTAATAAGTCAATCAGTATTCACATCTTTTACACTAAAAGCC 1004  
  
 RESULT 5  
 AAQ81005  
 ID AAQ81005 standard; DNA; 1032 BP.  
 XX  
 AC AAQ81005;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 22-AUG-1995 (first entry)  
 XX  
 DE BAGE tumor rejection antigen precursor DNA.  
 XX  
 KW BAGE; tumor rejection antigen precursor; diagnosis; HLA;  
 KW human leukocyte antigen MHC; major histocompatibility complex; TRAP;  
 KW cancer; melanoma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09500159-A1.  
 XX  
 PD 05-JAN-1995.  
 XX  
 PF 10-JUN-1994; 94WO-US006534.  
 XX  
 PR 17-JUN-1993; 93US-00079110.  
 PR 15-FEB-1994; 94US-00196630.  
 XX  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA Van Der Bruggen P, Boon-Falleur T, Coulie P, Renaud J;  
 PI WPI; 1995-051741/07.  
 XX  
 DR Nucleic acid coding for a tumour rejection antigen precursor - used to  
 PT develop prods. for the diagnosis and therapy of cancers, partic.  
 PT melanomas.  
 XX  
 PS Claim 1; Page 18; 33pp; English.  
 XX  
 CC The DNA encoding the tumor rejection antigen precursor BAGE was isolated  
 CC from a cDNA library prepared from RNA from human melanoma cell line M22-  
 CC MEL. The sequence and its encoded protein and derived peptide may be used  
 CC in the diagnosis and therapy of cancers, e.g. melanomas. (Updated on 25-  
 CC MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 1032 BP; 283 A; 212 C; 254 G; 283 T; 0 U; 0 Other;

Query Match 95.4%; Score 958; DB 2; Length 1032;  
 Best Local Similarity 96.8%; Pred. No. 1.6e-238;  
 Matches 999; Conservative 0; Mismatches 5; Indels 28; Gaps 1;  
  
 QY 1 CGCCAAATTTAGGGTCTCCGGTATCTCCGGTCTGAGTCTCTGTTCCCGGCTTAGAGGACC 60  
 Db 1 CGCCAAATTTAGGGTCTCCGGTATCTCCGGTCTGAGTCTCTGTTCCCGGCTTAGAGGACC 60  
 QY 61 AGGAGAGGGGAGCTGAGGCTGAGGCTGTAAACCGTGGCTCTCACTCTGGATG 120  
 Db 61 AGGAGAGGGGAGCTGAGGCTGAGGCTGTAAACCGTGGCTCTCACTCTGGATG 120  
 QY 121 GTGTGGCAACAGAGATGCGAGCGAGCTGAGTGTAGGAGGGGGCTCAGCGGTAGG 180  
 Db 121 GTGTGGCAACAGAGATGCGAGCGAGCTGAGTGTAGGAGGGGGCTCAGCGGTAGG 180  
 QY 181 AGTGGGCTGAGGAGCTAAGATGCGGCGCAGAGCGGTTTTCTGGCAATGTCTGCCAGC 240  
 Db 181 AGTGGGCTGAGGAGCTAAGATGCGGCGCAGAGCGGTTTTCTGGCAATGTCTGCCAGC 240  
 QY 241 TGCTCCAAAGCCAGGCTGATGAGGAGGAGTCCCTGTGGTGTGAGGTTGAGGCTG 300  
 Db 241 TGCTCCAAAGCCAGGCTGATGAGGAGGAGTCCCTGTGGTGTGAGGTTGAGGCTG 300  
 QY 301 AAGCGGCACAGCTCTGTGCTTCTATCTTCTGAGGTTGTGGCAGCCACGGTGTATGGAGCG 360  
 Db 301 AAGCGGCACAGCTCTGTGCTTCTATCTTCTGAGGTTGTGGCAGCCACGGTGTATGGAGCG 360  
 QY 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGAGGTTTCACTGTGTACGCCAGGATGGTCTC 420  
 Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGAGGTTTCACTGTGTACGCCAGGATGGTCTC 420  
 QY 421 GATCTCTGAGCTCGTGTATCGCCCGCTTCCAAAGTCCCAAGTCCAGGATACAGCGAT 480  
 Db 421 GATCTCTGAGCTCGTGTATCGCCCGCTTCCAAAGTCCCAAGTCCAGGATACAGCGAT 480  
 QY 481 GTGCAATTTGTAAGCACTTTGGAGCCACTATCAATCTGTGAAGAGAAATGTACCCAGA 540  
 Db 481 GTGCAATTTGTAAGCACTTTGGAGCCACTATCAATCTGTGAAGAGAAATGTACCCAGA 540  
 QY 541 TGTATCATTTATCCCTGCTGCTGAGGAGCGGCTTCCAGGATTTCAAGTCCAGTCAATCTCC 600  
 Db 541 TGTATCATTTATCCCTGCTGCTGAGGAGCGGCTTCCAGGATTTCAAGTCCAGTCAATCTCC 600  
 QY 601 TGCTTTGTCCAGAACACATTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660  
 Db 601 TGCTTTGTCCAGAACACATTGACCAAGCTCTGAAAGATGTAAGTTTACTACGCATAGAC 660  
 QY 661 TTTTAAACTTCAACCAATGATTTACTGAAAAATAAACHAATGTTGTAAATTCCTGAGTGT 720  
 Db 661 TTTTAAACTTCAACCAATGATTTACTGAAAAATAAACHAATGTTGTAAATTCCTGAGTGT 720  
 QY 721 TATTCTACTTGTATTTAAAGTAATAATACATATCATTAATCAATTAATCTGAGGATCATTTGCC 780  
 Db 721 TATTCTACTTGTATTTAAAGTAATAATACATATCATTAATCAATTAATCTGAGGATCATTTGCC 780  
 QY 781 AGAGATTGTTGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAATPCCAAAAGTT 840  
 Db 781 AGAGATTGTTGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAATPCCAAAAGTT 840  
 QY 841 GTTTTATTTTAAACATTTTAAAAACCACTGTAGAATGATTAATAGGAGTGTGC 932  
 Db 841 GTTTTATTTTAAACATTTTAAAAACCACTGTAGAATGATTAATAGGAGTGTGC 932  
 QY 932 AGTATTTCTGACATATATCAATTAATTAATAAGTCAATCAGTATTAACATCTTTTA 992  
 Db 932 AGTATTTCTGACATATATCAATTAATTAATAAGTCAATCAGTATTAACATCTTTTA 992  
 QY 992 AGTATTTCTGACATATATCAATTAATTAATAAGTCAATCAGTATTAACATCTTTTA 1020  
 Db 992 AGTATTTCTGACATATATCAATTAATTAATAAGTCAATCAGTATTAACATCTTTTA 1020



Db 6578 TTTCAGCAATGTCATTTTGTAAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAA 6519  
 Qy 531 TGTACCCAGATGATCATATTCCTGTGCTGCAGGAGCCGGCTCCTTTCAGGATTTTCAGT 590  
 Db 6518 TGTACCCAGATGATCATATTCCTGTGCTGCAGGAGCCGGCTCCTTTCAGGATTTTCAGT 6459  
 Qy 591 CACATCTTCCTGCTTTCCTCAGAACACATTCGACCAAGCTCCTGAAAGATGTAGTTTACT 650  
 Db 6458 CACATCTTCCTGCTTTCCTCAGAACACATTCGACCAAGCTCCTGAAAGATGTAGTTTACT 6399  
 Qy 651 ACGCATAGACTTTTAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAAT 710  
 Db 6398 ACGCATAGACTTTTAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAAT 6339  
 Qy 711 CCCTGAGTGATTTTCTACTGTTTAAAGGTAATTAATACATTAATCAATTAATTTTACT 770  
 Db 6338 CCCTGAGTGATTTTCTACTGTTTAAAGGTAATTAATACATTAATCAATTAATTTTACT 6279  
 Qy 771 GATCAATGCCAGAGATTTGGGAGGGAAATGTTATCAACGGTTTCAATGAAATTTAAAT 830  
 Db 6278 GATCGTTGCCAGAGATTTGGGAGGGAAATGTTATCAACGGTTTCAATGAAATTTAAAT 6219  
 Qy 831 CCAAAAAGTTTATTCCTCAGAAAAATCAATTAAGTTTGGATGTTTATTTTATTTTAAAC 890  
 Db 6218 CCAAAAAGTTTATTCCTCAGAAAAATCAATTAAGTTTGGATGTTTATTTTATTTTAAAC 6159  
 Qy 891 ATTTTAAATTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 950  
 Db 6158 ATTTTAAATTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 6099  
 Qy 951 TATAAATTTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 1004  
 Db 6098 TATAAATTTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 6045

RESULT 8

AAK99658/c  
 ID AAK99658 standard; DNA; 11162 BP.  
 AC AAK99658;  
 XX AAK99658;  
 DT 08-JUL-2002 (first entry)  
 DE DNA of the gene vector sequence (pGEM).  
 KW Cytostatic; human; gene leading sequence; short arm D group; tumour;  
 KW gene therapy; short arm G group; genetic disease; gene vector sequence;  
 KW pGEM; ds.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200220803-A1.  
 XX WO200220803-A1.  
 PD 14-MAR-2002.  
 XX 16-FEB-2001; 2001WO-CN000126.  
 XX 17-JUL-2000; 2000WO-CN000203.  
 XX (XIA/J) XIA J.  
 PA XIA J;  
 XX XIA J;  
 XX WPI; 2002-304383/34.  
 XX Human source chromosome-originated leading sequence as target gene in  
 PT vector for expression, useful in gene therapy and production of protein  
 PT e.g. for treating tumor and genetic diseases.  
 XX Example 2; Page 77-82; 865p; Chinese.  
 PS The invention relates to a human source gene leading sequence selected  
 XX from a DNA sequence in the chromosomal short arm of D and G groups, or  
 CC

CC DNA sequences having not less than 50% homology with the above sequence,  
 CC and does not have any important physiological functions. The invention  
 CC also relates to a gene vector and gene expression strategy. The leading  
 CC sequence is useful in gene therapy and production of protein e.g. for  
 CC treating tumour and genetic diseases. This polynucleotide sequence  
 CC represents the DNA of the gene vector sequence (pGEM) of the invention  
 XX  
 SQ Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;  
 Query Match 52.2%; Score 524.4; DB 6; Length 11162;  
 Best Local Similarity 98.9%; Pred. No. 2.1e-125;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 471 TTACAGCAATGTCATTTTGTAAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAA 530  
 Db 6578 TTTCAGCAATGTCATTTTGTAAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGAGAA 6519  
 Qy 531 TGTACCCAGATGATCATATTCCTGTGCTGCAGGAGCCGGCTCCTTTCAGGATTTTCAGT 590  
 Db 6518 TGTACCCAGATGATCATATTCCTGTGCTGCAGGAGCCGGCTCCTTTCAGGATTTTCAGT 6459  
 Qy 591 CACATCTTCCTGCTTTCCTCAGAACACATTCGACCAAGCTCCTGAAAGATGTAGTTTACT 650  
 Db 6458 CACATCTTCCTGCTTTCCTCAGAACACATTCGACCAAGCTCCTGAAAGATGTAGTTTACT 6399  
 Qy 651 ACGCATAGACTTTTAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAAT 710  
 Db 6398 ACGCATAGACTTTTAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAAT 6339  
 Qy 711 CCCTGAGTGATTTTCTACTGTTTAAAGGTAATTAATACATTAATCAATTAATTTTACT 770  
 Db 6338 CCCTGAGTGATTTTCTACTGTTTAAAGGTAATTAATACATTAATCAATTAATTTTACT 6279  
 Qy 771 GATCAATGCCAGAGATTTGGGAGGGAAATGTTATCAACGGTTTCAATGAAATTTAAAT 830  
 Db 6278 GATCGTTGCCAGAGATTTGGGAGGGAAATGTTATCAACGGTTTCAATGAAATTTAAAT 6219  
 Qy 831 CCAAAAAGTTTATTCCTCAGAAAAATCAATTAAGTTTGGATGTTTATTTTATTTTAAAC 890  
 Db 6218 CCAAAAAGTTTATTCCTCAGAAAAATCAATTAAGTTTGGATGTTTATTTTATTTTAAAC 6159  
 Qy 891 ATTTTAAATTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 950  
 Db 6158 ATTTTAAATTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 6099  
 Qy 951 TATAAATTTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 1004  
 Db 6098 TATAAATTTTAAAGTCAATCAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT 6045  
 RESULT 9  
 ABA99937/c  
 ID ABA99937 standard; DNA; 11162 BP.  
 XX ABA99937;  
 AC ABA99937;  
 XX ABA99937;  
 DT 05-JUL-2002 (first entry)  
 XX Human TNK-TPA DNA.  
 DE Human; tissue-type plasminogen activator; TNK-TPA; thrombolytic;  
 KW thrombosis; thrombus; blood vessel; ds.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200222832-A1.  
 XX WO200222832-A1.  
 PD 21-MAR-2002.  
 XX 16-FEB-2001; 2001WO-CN000127.  
 XX 04-SEP-2000; 2000WO-CN000260.  
 XX

PA (XIAJ/) XIA J.  
 XX Xia J;  
 XX WPI; 2002-362350/39.  
 XX Cell line expressing mutant human tissue-type plasminogen activator, for  
 PT use e.g. in treating thrombosis by local dissolution of thrombus in blood  
 PT vessels.  
 XX  
 XX  
 XX Claim 2; Page 20-25; 29pp; Chinese.  
 XX This invention describes a novel cell line expressing mutant human tissue  
 CC -type plasminogen activator having an Accession Number of C200006  
 CC and which has thrombolytic activity. The product of the invention can be  
 CC used for treating thrombosis by local dissolution of thrombus in blood  
 CC vessels. This sequence encodes a human tissue-type plasminogen activator  
 CC TNK-TFA described in the disclosure of the invention  
 XX  
 XX Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;  
 SQ  
 Query Match 52.2%; Score 524.4; DB 6; Length 11162;  
 Best Local Similarity 98.9%; Pred. No. 2.1e-125;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 471 TTACGCGATGTCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
 Db 6578 TTTGAGCAATGTGCAATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 6519  
 Qy 531 TGTACCCAGATGTATCATTTCTTGTGCTGAGGAGCCGCTCTTTCAGGATTTTCAGT 590  
 Db 6518 TGTACCCAGATGTATCATTTCTTGTGCTGAGGAGCCGCTCTTTCAGGATTTTCAGT 6459  
 Qy 591 CACATCTTCCTGCTTTGTCAGAACACATTCACCAAGCTCTGAAAGATGTAAGTTTACT 650  
 Db 6458 CACATCTTCCTGCTTTGTCAGAACACATTCACCAAGCTCTGAAAGATGTAAGTTTACT 6399  
 Qy 651 ACGCATAGACTTTTAAACTTCAACCAATGATTTACTGAAATAACAAATGTTGTAAAT 710  
 Db 6398 ACGCATAGACTTTTAAACTTCAACCAATGATTTACTGAAATAACAAATGTTGTAAAT 6339  
 Qy 711 CCTGAGTGTATTCTACTTGTATTTAAAGGTAATATACATATCATTTAAATCTGAGG 770  
 Db 6338 CCTGAGTGTATTCTACTTGTATTTAAAGGTAATATACATATCATTTAAATCTGAGG 6279  
 Qy 771 GATCATTTGCCAGAGATTTGGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAAT 830  
 Db 6278 GATCATTTGCCAGAGATTTGGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAAT 6219  
 Qy 831 CCAAAAAGTTATTTCCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATCTTAAAC 890  
 Db 6218 CCAAAAAGTTATTTCCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATCTTAAAC 6159  
 Qy 891 ATTTTAAACCACTGTAGATGATGTAATAGGAGCTGCGAGTATTTCTGACATATAC 950  
 Db 6158 ATTTTAAACCACTGTAGATGATGTAATAGGAGCTGCGAGTATTTCTGACATATAC 6099  
 Qy 951 TATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTACACTTAAAGGCC 1004  
 Db 6098 TATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTACACTTAAAGGCC 6045  
 RESULT 10  
 AAH77500/C  
 ID AAH77500 standard; DNA; 13928 BP.  
 AC AAH77500;  
 XX  
 XX 04-DEC-2001 (first entry)  
 XX Haemophilia B genetic therapy related oligonucleotide #3.  
 XX  
 XX Haemophilia B; coagulant; ds.

XX Unidentified.  
 XX CN1302664-A.  
 XX 11-JUL-2001.  
 PD  
 XX 19-JAN-2001; 2001CN-00102830.  
 XX  
 XX 30-AUG-2000; 2000CN-00113652.  
 XX  
 XX (XIAJ/) XIA J.  
 XX Xia J;  
 XX WPI; 2001-550515/62.  
 XX Gene medicine, useful for treating hemophilia B, comprises the first gene  
 PT carrier-FIX recombination body and the DNA sequence which is 50% or more  
 PT homogenous with first gene sequence as the pilot sequence of therapeutic  
 PT gene.  
 XX  
 XX Disclosure; Page 19-25(Disclosure); 29pp; Chinese.  
 XX The present invention relates to a genetic therapy for treating  
 CC haemophilia B, which contains the gene carrier-FIX recombination body,  
 CC which uses the DNA sequence of the gene without important physiologic  
 CC function correlation on the short arm of a human chromosome in group D or  
 CC G and the DNA sequence which is 50% or more homogenous with the above  
 CC gene sequence as the pilot sequence of the therapeutic gene. This is  
 CC useful in the treatment of haemophilia B. The present sequence is an  
 CC oligonucleotide described in the exemplification of the invention  
 XX  
 XX Sequence 13928 BP; 3744 A; 3345 C; 3126 G; 3713 T; 0 U; 0 Other;  
 SQ  
 Query Match 52.2%; Score 524.4; DB 4; Length 13928;  
 Best Local Similarity 98.9%; Pred. No. 2.3e-125;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 471 TTACGCGATGTCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
 Db 9344 TTTGAGCAATGTGCAATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 9285  
 Qy 531 TGTACCCAGATGTATCATTTCTTGTGCTGAGGAGCCGCTCTTTCAGGATTTTCAGT 590  
 Db 9284 TGTACCCAGATGTATCATTTCTTGTGCTGAGGAGCCGCTCTTTCAGGATTTTCAGT 9225  
 Qy 591 CACATCTTCCTGCTTTGTCAGAACACATTCACCAAGCTCTGAAAGATGTAAGTTTACT 650  
 Db 9224 CACATCTTCCTGCTTTGTCAGAACACATTCACCAAGCTCTGAAAGATGTAAGTTTACT 9165  
 Qy 651 ACGCATAGACTTTTAAACTTCAACCAATGATTTACTGAAATAACAAATGTTGTAAAT 710  
 Db 9164 ACGCATAGACTTTTAAACTTCAACCAATGATTTACTGAAATAACAAATGTTGTAAAT 9105  
 Qy 711 CCTGAGTGTATTCTACTTGTATTTAAAGGTAATATACATATCATTTAAATCTGAGG 770  
 Db 9104 CCTGAGTGTATTCTACTTGTATTTAAAGGTAATATACATATCATTTAAATCTGAGG 9045  
 Qy 771 GATCATTTGCCAGAGATTTGGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAAT 830  
 Db 9044 GATCATTTGCCAGAGATTTGGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAAT 8985  
 Qy 831 CCAAAAAGTTATTTCCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATCTTAAAC 890  
 Db 8984 CCAAAAAGTTATTTCCTCAGAAAAATCAAAATAAGTTTGCATGTTTTTATCTTAAAC 8925  
 Qy 891 ATTTTAAACCACTGTAGATGATGTAATAGGAGCTGTCAGTATTTCTGACATATAC 950  
 Db 8924 ATTTTAAACCACTGTAGATGATGTAATAGGAGCTGTCAGTATTTCTGACATATAC 8865  
 Qy 951 TATAAAATTTATAAAAGTCAATCAGTATTCACATCTTTTACACTTAAAGGCC 1004

Db 8864 TATAAAATTATTAATAAGTCATCAGTATTCAACATCTTTTACACTAAAAAGCC 8811

RESULT 11  
AAK99657/c  
ID AAK99657 standard; DNA; 107602 BP.  
XX AC  
XX AC AAK99657;  
XX  
DT 08-JUL-2002 (first entry)  
XX  
DE DNA of the PAC clone sequence.  
XX  
KW Cytostatic; human; gene leading sequence; short arm D group; tumour;  
KW Gene therapy; short arm G group; genetic disease; PAC clone sequence; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200220803-A1.  
XX  
PD 14-MAR-2002.  
XX  
PF 16-FEB-2001; 2001WO-CN000126.  
XX  
PR 17-JUL-2000; 2000WO-CN000203.  
XX  
PA (XIAJ/) XIA J.  
XX  
PI Xia J;  
XX  
DR WPI; 2002-304383/34.  
XX  
XX Human source chromosome-originated leading sequence as target gene in  
PT vector for expression, useful in gene therapy and production of protein  
PT e.g. for treating tumor and genetic diseases.  
XX  
PS Example 1; Page 28-77; 86pp; Chinese.  
XX  
CC The invention relates to a human source gene leading sequence selected  
CC from a DNA sequence in the chromosomal short arm of D and G groups, or  
CC DNA sequences having not less than 50% homology with the above sequence,  
CC and does not have any important physiological functions. The invention  
CC also relates to a gene vector and gene expression strategy. The leading  
CC sequence is useful in gene therapy and production of protein e.g. for  
CC treating tumour and genetic diseases. This polynucleotide sequence  
CC represents the DNA of the PAC clone sequence of the invention  
XX  
SQ Sequence 107602 BP; 36346 A; 19466 C; 18991 G; 32779 T; 0 U; 20 Other;  
Query Match 52.2%; Score 524.4; DB 6; Length 107602;  
Best Local Similarity 98.9%; Pred. No. 5.8e-125;  
Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 471 TTACAGCGATGTCATTTGTAGCAGCTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
Db 77748 TTTGAGCAATGTGCATTTGTAGCAGCTTTGGAGCCACTATCAATGCTGTGAAGAGAA 77689

QY 531 TGTACCCAGATGATCATTTATCCTTGTGCTGAGGAGCCGGCTCTTTCCAGGATTTCACT 590  
Db 77688 TGTACCCAGATGATCATTTATCCTTGTGCTGAGGAGCCGGCTCTTTCCAGGATTTCACT 77629

QY 591 CACATCTTCTGCTTTGTCAGAGAACATTTGACCAAGCTCTTGAAGAGATGTAAGTTTACT 650  
Db 77628 CACATCTTCTGCTTTGTCAGAGAACATTTGACCAAGCTCTTGAAGAGATGTAAGTTTACT 77569

QY 651 ACCCATAGACTTTTAACTTCAACCAATGATTTACTGAAATAACAAATGTTGTAAT 710  
Db 77568 ACGGATAGACTTTTAACTTCAACCAATGATTTACTGAAATAACAAATGTTGTAAT 77509

QY 711 CCTGAGTGTTATCTTACTGTTATTAAGAGGTAATAATACATAATCATTTAAATCTGAGG 770  
Db 77508 CCTGAGTGTTATCTTACTGTTATTAAGAGGTAATAATACATAATCATTTAAATCTGAGG 77449

QY 771 GATCATTTGCCAGAGATTTGTTGGGAGGGAATTTATCAACGGTTTCATTGAAATTAAT 830  
Db 77448 GATCGTTGCCAGAGATTTGTTGGGAGGGAATTTATCAACGGTTTCATTGAAATTAAT 77389

QY 831 CCAGAAAAGTTATTTCTCTCAGAAAAATCAAAATAAGTTTTCATGTTTATTTTATTTTAAAC 890  
Db 77388 CCAGAAAAGTTATTTCTCTCAGAAAAATCAAAATAAGTTTTCATGTTTATTTTATTTTAAAC 77329

QY 891 ATTTTAAAAACCACTGTAGATGATGTAATAGGAGCTGTGAGTATTTCTGACATATAC 950  
Db 77328 ATTTTAAAAACCACTGTAGATGATGTAATAGGAGCTGTGAGTATTTCTGACATATAC 77269

QY 951 TATAAAATTTAAAAAGTCAATCAGTATTCAGATTCCTTTTACACTAAAAAGCC 1004  
Db 77268 TATAAAATTTAAAAAGTCAATCAGTATTCAGATTCCTTTTACACTAAAAAGCC 77215

RESULT 12  
ABL54503/c  
ID ABL54503 standard; DNA; 107612 BP.  
XX  
AC ABL54503;  
XX  
DT 27-MAY-2002 (first entry)  
XX  
DE Human PAC clone sequence SEQ ID NO 1.  
XX  
KW Human; D chromosome; G chromosome; gene therapy; nucleolus; TPA; FIX;  
KW thrombotic disease; haemophilia; PAC clone; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200210375-A1.  
XX  
PD 07-FEB-2002.  
XX  
PF 17-JUL-2000; 2000WO-CN000203.  
XX  
PR 17-JUL-2000; 2000WO-CN000203.  
XX  
PA (XIAJ/) XIA J.  
XX  
PI Xia J;  
XX  
DR WPI; 2002-291858/33.  
XX  
PT Human source chromosome-originated gene leading sequence of target gene  
PT in vector for transfer into nucleolus of cell for active expression,  
PT useful in gene therapy and production of desirable proteins.  
XX  
PS Disclosure; Page 20-69; 76pp; Chinese.  
XX  
CC The invention relates to a human source gene leading sequence selected  
CC from DNA sequences in the short arms of the human D or G group  
CC chromosome, which does not have any important physiological function. The  
CC leading sequence is useful in gene therapy for targeting a gene in a  
CC vector for transfer into nucleolus of cells for active expression. The  
CC sequence may be used for the production of desirable proteins for  
CC treating, e.g. thrombotic diseases and haemophilia B with TPA and FIX,  
CC respectively. DNA sequences such as those in the short arms of the human  
CC D or G group chromosome do not have any important physiological function  
CC and are safely used in gene therapy. The present sequence is that of a  
CC human PAC clone sequence, useful to the invention  
XX  
SQ Sequence 107612 BP; 36302 A; 19500 C; 19011 G; 32769 T; 0 U; 30 Other;

Query Match 52.2%; Score 524.4; DB 6; Length 107612;  
Best Local Similarity 98.9%; Pred. No. 5.8e-125;  
Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 471 TTACAGCGATGTCATTTGTAGCAGCTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
Db 77749 TTTGAGCAATGTGCATTTGTAGCAGCTTTGGAGCCACTATCAATGCTGTGAAGAGAA 77690

CC claims the KGF PCR primers AAC87589 and AAC87590 for use in detecting all or part of the KGF gene. The nucleic acid sequences can be used to detect abnormalities in chromosomal region 9p11 and thus give an indication of an individual's risk of developing a 9p11-associated condition. The present sequence represents human chromosomal region 9q11

CC Sequence 57728 BP; 16960 A; 12164 C; 11741 G; 16863 T; 0 U; 0 Other;

Query Match 18.1%; Score 181.8; DB 4; Length 57728;

Best Local Similarity 89.9%; Pred. No. 4.4e-36;

Matches 195; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CGCCAAATTTAGGGTCTCCGGTATCTCCCGTGTAGCTCTCTGTTCCCGGCTTAGAGGACC 60

DB 45622 CGCCAGTTTAGGCTCTCCAGTAICTCTCTCCGAGCTCTCTGTTCCCGGCTTAGAGGACC 45681

QY 61 AGGAGAGGGGGAGCTGGAGGCTGGAGCTGTAAACACCATGGCTCGTCTCGCTCTGGATG 120

DB 45682 AGGAGAGGGGGAGCTGGAGGCTGGAGCTGTAAACACCATGGCTCGTCTCGCTCTGGATG 45741

QY 121 GTGGTGTCAACAGAGATGGCAGCGCTGGAGTGTAGGAGGGCGGCTTAGAGGCTTAGG 180

DB 45742 GTGGTGTCAACAGAGATGGCAGCGCTGGAGTGTAGGAGGGCGGCTTAGAGGCTTAGG 45801

QY 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGT 217

DB 45802 AGGTAGTCTGAGCAGTAGGAGGGGGGCTGGAGCAGT 45838

RESULT 14

AAS46735

ID AAS46735 standard; DNA; 6292 BP.

XX AC AAS46735;

XX DT 18-DEC-2001 (first entry)

XX DE Tumour suppressor gene derived chemically modified sequence #459.

XX KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;

XX KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;

XX KW cytosine methylation; ds.

XX OS Homo sapiens.

XX PN WO200168912-A2.

XX PD 20-SEP-2001.

XX PF 15-MAR-2001; 2001WO-EP002955.

XX PR 15-MAR-2000; 2000DE-01013847.

XX PR 06-APR-2000; 2000DE-01019058.

XX PR 07-APR-2000; 2000DE-01019173.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-602752/68.

XX PT Fragments of chemically modified genes associated with tumor suppressor genes and oncogenes, useful in designing primers and probes for analyzing diseases associated with cytosine methylation state e.g. cancer.

XX PS Claim 1; SEQ ID NO 459; 27pp; English.

XX CC The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and

QY 531 TGTACCCAGATGATCATTTATCTCTGTCTGTCAGGAGCCGGCTCTCTTCAGGATTTCAAT 590

DB 77689 TGTACCCAGATGATCATTTATCTCTGTCTGTCAGGAGCCGGCTCTCTTCAGGATTTCAAT 77630

QY 591 CACATCTCTCTCTTGTCTGTCAGGAGCCGGCTCTCTTCAGGATTTCAAT 650

DB 77629 CACATCTCTCTCTTGTCTGTCAGGAGCCGGCTCTCTTCAGGATTTCAAT 77570

QY 651 ACCGATAGACTTTTAACTTCAACCAATGATTTACTGAAAATTAACAAATGTTTAAAT 710

DB 77569 ACCGATAGACTTTTAACTTCAACCAATGATTTACTGAAAATTAACAAATGTTTAAAT 77510

QY 711 CCTGAGTGTATCTACTTGTATTTAAAGGTTAATATACATTAATTAATCTGAGG 770

DB 77509 CCTGAGTGTATCTACTTGTATTTAAAGGTTAATATACATTAATTAATCTGAGG 77450

QY 771 GATCATTTGCCAGAGATTTTGGGGAGGAAATGTTATCAACGTTTTCATTGAAATTAAT 830

DB 77449 GATCATTTGCCAGAGATTTTGGGGAGGAAATGTTATCAACGTTTTCATTGAAATTAAT 77390

QY 831 CCAAAAGTTTATTTCTCAGAAAAATCAAAATAAAGTTTGGATTTTATTTCTTAAAC 890

DB 77389 CCAAAAGTTTATTTCTCAGAAAAATCAAAATAAAGTTTGGATTTTATTTCTTAAAC 77330

QY 891 ATTTTAAACCACTGTAGATGATGTAATAGGAGCTGTGCAATTTTCTGACATATAC 950

DB 77329 ATTTTAAACCACTGTAGATGATGTAATAGGAGCTGTGCAATTTTCTGACATATAC 77270

QY 951 TATAAATTTATAAAGTCAATCAGTATTCAACATCTTTTACATTAATAAGCC 1004

DB 77269 TATAAATTTATAAAGTCAATCAGTATTCAACATCTTTTACATTAATAAGCC 77216

RESULT 13

AAC87588

ID AAC87588 standard; DNA; 57728 BP.

XX AC AAC87588;

XX DT 16-MAR-2001 (first entry)

XX DE Human 9p11 chromosomal region.

XX KW Human; chromosome 9p11; abnormality; cancer; miscarriage;

XX KW spontaneous abortion; genetic susceptibility; diagnosis; KGF;

XX KW keratinocyte growth factor; ds.

XX OS Homo sapiens.

XX PN JP2000287684-A.

XX PD 17-OCT-2000.

XX PF 31-JAN-2000; 2000JP-00022688.

XX PR 05-FEB-1999; 99JP-00028705.

XX PA (ASAK) ASAHI BREWERIES LTD.

XX DR WPI; 2001-065570/08.

XX PT The base sequence of 9p11 chromosomal region participating to cancer and abortion.

XX PS Claim 1; Page 6-87; 88pp; Japanese.

XX CC The invention relates to human chromosomal region 9q11 (AAC87588).

XX CC Abnormalities in this region of the short arm of chromosome 9 is thought to be associated with miscarriage and cancer, as an ovarian cancer patient with a history of miscarriage was found to have a chromosomal inversion inv(9)(p11;q13). The 9p11 region contains the gene encoding keratinocyte growth factor (KGF), and the invention also specifically

CC 500 are missing from the sequence listing) sequences (ss) and sequences  
 CC complementary to (ss). The nucleic acid may be a peptide nucleic acid-  
 CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of  
 CC probes for detecting the cytosine methylation state and/or single  
 CC nucleotide polymorphisms and also to be used in an array for analysing  
 CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The  
 CC probes can also be used in a method for ascertaining genetic and/or  
 CC epigenetic parameters for the diagnosis and/or therapy of existing  
 CC diseases or the predisposition to specific diseases, by analysing  
 CC cytosine methylations. The parameters may be compared to another set of  
 CC genetic and/or epigenetic parameters, the differences serving as basis  
 CC for diagnosis and/or prognosis events which are disadvantageous to  
 CC patients. The present sequence is one of the 533 genomic sequences  
 CC derived from tumour suppressor genes and oncogenes. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 SQ Sequence 6292 BP; 1736 A; 99 C; 1304 G; 3153 T; 0 U; 0 Other;

Query Match 14.9%; Score 149.8; DB 4; Length 6292;  
 Best Local Similarity 80.6%; Pred. No. 3.3e-28;  
 Matches 175; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
 QY 1 CGCCAAATTTAGGGTCTCCGGTATCTCCCGTGGAGCTCTGTTCCTCCGGCTTAGAGGACC 60  
 Db 5001 CATTAGTTTAGGGTTTGGTATTTTCGTGAGTTGTTTTCGGTTTAGAGGATT 5060  
 QY 61 AGGAGAAAGGGGAGCTCGAGCTGGAGCCTGTAAACACCGTGGCTCTGCTCACTCTGGATG 120  
 Db 5061 AGGAGAAAGGGGAGTTGGAGTTGGAGTTTGTATATCGTGGTTCGTTTCGTTTGGATG 5120  
 QY 121 GTGGTGGACACAGAGTCGAGCGCAGCTGGAGTGTAGGAGGCGGCTGACCGGTAGG 180  
 Db 5121 GTGGTGGTATAGAGATGGTAGCGTAGTTGGAGTGTAGGAGGCGGCTTGGACCGGTAGG 5180  
 QY 181 AGTGGGCTGGAGCAGTAAGATGCGGCCACAGAGCGGT 217  
 Db 5181 AGTGGGCTGGAGTAGTAGTAAGATGCGGCTCGAGTAGT 5217

RESULT 15  
 AAC19397  
 ID AAC19397 standard; cDNA; 251 BP.  
 XX  
 AC AAC19397;  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX  
 DE Human secreted protein 5' EST, SEQ ID NO: 23472.  
 XX  
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1033401-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 21-FEB-2000; 2000EP-00200610.  
 XX  
 PR 26-FEB-1999; 99US-0122487P.  
 XX  
 PA (GIST ) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Duclert A, Giordano J;  
 XX  
 DR WPI; 2000-500381/45.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX  
 PS  
 XX Claim 1; SEQ ID NO 23472; 71pp + Sequence Listing; English.  
 CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
 CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
 CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors  
 CC  
 XX  
 SQ Sequence 251 BP; 49 A; 50 C; 99 G; 50 T; 0 U; 3 Other;  
 Query Match 10.6%; Score 106.2; DB 3; Length 251;  
 Best Local Similarity 88.8%; Pred. No. 1.6e-17;  
 Matches 111; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 QY 261 AAGGAGGAGTCCCTGTGCTGAGCTGGAGCTGGAGCTGAAGACGGCAGCTCTGTGC 320  
 Db 1 AAGGAGSATTCCCTGTGCTGAGCTKGWGGTTGTAGCCTGAAGATGGCAGCTCTGTGA 60  
 QY 321 TTTCATCTTCTGAGGTTGTGGCAGCCACCGTGTATGAGAGCGCAGCTCAACAGGAGCAATA 380  
 Db 61 TTTCATCTCTCGGTTGTGGCGCCACCGTGTATGAGAGCTCAACAGGAGTGGA 120  
 QY 381 GGAGG 385  
 Db 121 GGAGG 125  
 Search completed: July 9, 2004, 07:09:19  
 Job time : 641 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 05:43:15 ; Search time 105 Seconds  
(without alignments)  
5306.388 Million cell updates/sec

Title: US-10-081-108-1  
Perfect score: 1004  
Sequence: 1 CGCAATTAGGCTCCGG.....ATCTTTTACACTAAAAGCC 1004

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCUS COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004	100.0	1004	2	US-08-389-360-1
2	1004	100.0	1004	3	US-09-038-328-1
3	1004	100.0	1004	3	US-09-056-105-20
4	1004	100.0	1004	4	US-09-435-524-1
5	1004	100.0	1004	4	US-09-382-497-1
6	966	96.2	1032	1	US-08-079-110A-1
7	966	96.2	1032	1	US-08-196-630A-1
8	966	96.2	1032	3	US-08-573-186-1
9	83	8.3	392000	4	US-10-027-983-11
10	82.4	8.2	859	4	US-09-535-008-58
11	82.4	8.2	4698	4	US-09-439-261-34
12	82.4	8.2	4698	4	US-09-427-613-33
13	82.4	8.2	15602	4	US-09-844-634-17
14	82.4	8.2	35100	4	US-08-306-691B-19
15	82.4	8.2	35100	5	PCT-US93-06251-19
16	82.4	8.2	39982	4	US-09-820-924-3
17	82.4	8.2	84495	4	US-09-797-906-3
18	82.4	8.2	319608	4	US-09-539-3330-1
19	82.4	8.2	319608	4	US-09-679-409-1
20	82.2	8.2	7210	2	US-08-257-963B-10
21	82.2	8.2	7210	4	US-08-367-841A-10
22	82.2	8.2	7210	5	PCT-US95-07201-10
23	82.2	8.2	14581	4	US-08-520-3730-4
24	82.2	8.2	22481	4	US-08-367-841A-43
25	82.2	8.2	22481	5	PCT-US95-07201-43
26	82.2	8.2	22484	4	US-09-875-223-2
27	82.2	8.2	22484	4	US-09-875-114-2

28	81.8	8.1	453	4	US-09-621-976-12690	Sequence 12690, A
C 29	81.8	8.1	1268	4	US-09-369-247-42	Sequence 42, Appl
C 30	81.8	8.1	6804	4	US-09-740-041-3	Sequence 3, Appl
C 31	81.4	8.1	311	4	US-09-621-976-2093	Sequence 2093, Ap
32	81	8.1	7313	4	US-09-620-312D-259	Sequence 259, App
33	80.8	8.0	460	4	US-09-918-686-26	Sequence 26, Appl
C 34	80.8	8.0	954	4	US-09-621-976-1616	Sequence 1616, Ap
35	80.8	8.0	1001	4	US-09-641-638-376	Sequence 376, App
36	80.8	8.0	1744	4	US-09-511-625B-19	Sequence 19, Appl
37	80.8	8.0	2280	4	US-09-702-705-321	Sequence 321, App
38	80.8	8.0	2280	4	US-09-736-457-321	Sequence 321, App
39	80.8	8.0	2280	4	US-09-614-124B-321	Sequence 321, App
40	80.8	8.0	2280	4	US-09-671-325-321	Sequence 321, App
41	80.8	8.0	2280	4	US-09-589-184-321	Sequence 321, App
42	80.8	8.0	17425	4	US-09-511-625B-5	Sequence 5, Appl
C 43	80.8	8.0	18736	4	US-09-740-035-3	Sequence 3, Appl
44	80.8	8.0	20674	4	US-09-641-638-651	Sequence 651, App
C 45	80.8	8.0	29629	4	US-09-729-995-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-08-389-360-1  
; Sequence 1, Application US/08389360  
; Patent No. 5877017  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen et al.  
; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
; TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/389,360  
; FILING DATE: Herewith  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/196,630  
; FILING DATE: February 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/079,110  
; FILING DATE: June 17, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pasqualini, Patricia A.  
; REGISTRATION NUMBER: 34,894  
; REFERENCE/DOCKET NUMBER: LUD 5310.2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1004  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-389-360-1  
Query Match 100.0%; Score 1004; DB 2; Length 1004;  
Best Local Similarity 100.0%; Pred. No. 3.7e-258;  
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	C	G	C	C	A	A	T	T	T	A	G	G	T	C	T	C	C	G	T	A	T	C	C	C	G	T	G	A	G	T	G	T	C	T	G	T	T	C	C	G	G	T	T	A	G	A	G	A	C	C	60
Db	1	C	G	C	C	A	A	T	T	T	A	G	G	T	C	T	C	C	G	T	A	T	C	C	C	G	T	G	A	G	T	G	T	C	T	G	T	T	C	C	G	G	T	T	A	G	A	G	A	C	C	60
QY	61	A	G	G	A	A	G	G	G	G	G	A	G	T	G	G	A	G	G	C	T	G	T	A	A	C	C	G	T	G	G	T	G	T	C	T	C	A	C	T	C	T	G	T	C	T	G	A	G	A	C	120
Db	61	A	G	G	A	A	G	G	G	G	A	G	T	G	G	A	G	G	C	T	G	T	A	A	C	C	G	T	G	G	T	G	T	C	T	C	A	C	T	C	T	G	T	C	T	G	A	G	A	C	120	
QY	121	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	T	A	G	A	G	G	G	C	G	C	T	T	A	G	A	G	C	G	T	A	G	180						
Db	121	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	T	A	G	A	G	G	C	G	C	T	T	A	G	A	G	C	G	T	A	G	180							
QY	181	A	G	T	G	G	G	T	G	G	A	G	T	A	A	G	T	G	G	C	C	A	G	G	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	240			
Db	181	A	G	T	G	G	G	T	G	G	A	G	T	A	A	G	T	G	G	C	C	A	G	G	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	240				
QY	241	T	G	T	C	C	A	A	G	C	C	A	G	T	A	A	G	A	G	A	G	T	C	C	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	300			
Db	241	T	G	T	C	C	A	A	G	C	C	A	G	T	A	A	G	A	G	A	G	T	C	C	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	300			
QY	301	A	A	G	A	C	G	C	A	C	A	G	T	C	T	G	T	C	T	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	360				
Db	301	A	A	G	A	C	G	C	A	C	A	G	T	C	T	G	T	C	T	C	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	360				
QY	361	G	C	A	G	C	T	C	A	A	G	A	G	A	G	A	G	A	G	A	G	T	C	A	C	T	T	C	A	C	T	G	T	G	T	C	A	C	T	G	T	C	A	C	T	G	T	C	420			
Db	361	G	C	A	G	C	T	C	A	A	G	A	G	A	G	A	G	A	G	A	G	T	C	A	C	T	T	C	A	C	T	G	T	C	A	C	T	G	T	C	A	C	T	G	T	C	420					
QY	421	G	A	T	C	T	C	C	T	C	A	C	T	C	G	T	G	T	C	C	T	T	C	A	A	A	G	T	G	C	C	G	A	T	T	A	C	A	G	T	A	C	A	G	T	A	C	480				
Db	421	G	A	T	C	T	C	C	T	C	A	C	T	C	G	T	G	T	C	C	T	T	C	A	A	A	G	T	G	C	C	G																				

Patent No. 6110694  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen et al.  
TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pelfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/038,328  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/389,360  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/079,110  
FILING DATE: June 17, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: LUD 5310.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1004  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-038-328-1

RESULT 2  
US-09-038-328-1  
; Sequence 1, Application US/09038328

Db 361 GCAGCTCAACAGGAGCAATAGGAGAGATGGAGTTTCACTGTGTGACCCAGGATGTCCTC 420  
Qy 421 GATCTCCTGACCTCGTGTATCGCGCCGCTTGGCCCTTCCAAAGTGGCGAGATTACAGCAT 480  
Db 421 GATCTCCTGACCTCGTGTATCGCGCCGCTTGGCCCTTCCAAAGTGGCGAGATTACAGCAT 480  
Qy 481 GTGCATTTTGTAGACACTTTGGAGCCACTATCAAAATGCTGTGAAGAAATGTACCCAGA 540  
Db 481 GTGCATTTTGTAGACACTTTGGAGCCACTATCAAAATGCTGTGAAGAAATGTACCCAGA 540  
Qy 541 TGTATCAATATCTTGTGTGAGGAGCCGCTTTCAGGATTTTCAGTCAATCTTCC 600  
Db 541 TGTATCAATATCTTGTGTGAGGAGCCGCTTTCAGGATTTTCAGTCAATCTTCC 600  
Qy 601 TGTCTTGTCCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGGCATAGAC 660  
Db 601 TGTCTTGTCCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGGCATAGAC 660  
Qy 661 TTTTAACTTCAACCAATGATTTACTGAAATTAACAAATGTTGTAATTCCTGAGTGT 720  
Db 661 TTTTAACTTCAACCAATGATTTACTGAAATTAACAAATGTTGTAATTCCTGAGTGT 720  
Qy 721 TATCTACTTGTATTAAGGTAATTAATCAATCAATTAATAATCTGAGGATCAATTGCC 780  
Db 721 TATCTACTTGTATTAAGGTAATTAATCAATCAATTAATAATCTGAGGATCAATTGCC 780  
Qy 781 AGAGATTTGGGAGGGAATGTTATCAACGGTTTCATGAAATTAATTAATAAGTT 840  
Db 781 AGAGATTTGGGAGGGAATGTTATCAACGGTTTCATGAAATTAATTAATAAGTT 840  
Qy 841 ATTTCTTCAGAAAAATCAATAAGTTTGCATGTTTTTATCTTAAACATTTTAAAA 900  
Db 841 ATTTCTTCAGAAAAATCAATAAGTTTGCATGTTTTTATCTTAAACATTTTAAAA 900  
Qy 901 CCAGTGTAGATGATGTAATAGGAGCTGTGCAATTTTCTGACATATATATAAATTA 960  
Db 901 CCAGTGTAGATGATGTAATAGGAGCTGTGCAATTTTCTGACATATATATAAATTA 960  
Qy 961 TTAAGAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004  
Db 961 TTAAGAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004

RESULT 3

US-09-056-105-20  
; Sequence 20, Application US/09056105  
; Patent No. 6287569

GENERAL INFORMATION:  
; APPLICANT: KIPPS, THOMAS J.  
; APPLICANT: WU, YUNQI

; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR  
; FILE REFERENCE: 233/221

; CURRENT APPLICATION NUMBER: US/09/056,105  
; EARLIER FILING DATE: 1998-04-06

; EARLIER APPLICATION NUMBER: 60/043,467  
; EARLIER FILING DATE: 1997-04-10

; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 20  
; LENGTH: 1004

; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-056-105-20

Query Match 100.0%; Score 1004; DB 3; Length 1004;

Best Local Similarity 100.0%; Pred. No. 3,7e-258; Indels 0; Gaps 0;  
Matches 1004; Conservative 0; Mismatches 0

Qy 1 CGCAATTTAGGCTCTCCGGTATCTCCCGTACGCTGCTCTGTCTCCGGCTTAGAGACC 60

Db 1 CGCAATTTAGGCTCTCCGGTATCTCCCGTACGCTGCTCTGTCTCCGGCTTAGAGACC 60

Qy 61 AGGAGAGGGGAGCTGGAGGCTGTAAACACCGTGGCTCTCTCACTCTGATG 120  
Db 61 AGGAGAGGGGAGCTGGAGGCTGTAAACACCGTGGCTCTCTCACTCTGATG 120  
Qy 121 GTGTGGCAACAGAGATGGCAGCGAGTGTAGGAGGGGGCTGAGCGGTAGG 180  
Db 121 GTGTGGCAACAGAGATGGCAGCGAGTGTAGGAGGGGGCTGAGCGGTAGG 180  
Qy 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGTCATGTCGCCAGC 240  
Db 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGTCATGTCGCCAGC 240  
Qy 241 TGTCTCAAGCCAGGCTGATGAAGAGGAGTCCCTGTGGTGTGAGCTGGAAGTGTGAGCCTG 300  
Db 241 TGTCTCAAGCCAGGCTGATGAAGAGGAGTCCCTGTGGTGTGAGCTGGAAGTGTGAGCCTG 300  
Qy 301 AAGAGGCGACAGCTGTGCTTCACTTCTGAGTGTGGCAGCGACCGTATGGAGACG 360  
Db 301 AAGAGGCGACAGCTGTGCTTCACTTCTGAGTGTGGCAGCGACCGTATGGAGACG 360  
Qy 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTCACCTGTGTCAGCCAGATGGTCTC 420  
Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTCACCTGTGTCAGCCAGATGGTCTC 420  
Qy 421 GATCTCCTGACCTGTGTATCGCGCCGCTTGGCCCTTCCAAAGTGGCGAGATTACAGCAT 480  
Db 421 GATCTCCTGACCTGTGTATCGCGCCGCTTGGCCCTTCCAAAGTGGCGAGATTACAGCAT 480  
Qy 481 GTGCATTTTGTAGACACTTTGGAGCCACTATCAAAATGCTGTGAAGAAATGTACCCAGA 540  
Db 481 GTGCATTTTGTAGACACTTTGGAGCCACTATCAAAATGCTGTGAAGAAATGTACCCAGA 540  
Qy 541 TGTATCAATATCTTGTGTGAGGAGCCGCTTTCAGGATTTTCAGTCAATCTTCC 600  
Db 541 TGTATCAATATCTTGTGTGAGGAGCCGCTTTCAGGATTTTCAGTCAATCTTCC 600  
Qy 601 TGTCTTGTCCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGGCATAGAC 660  
Db 601 TGTCTTGTCCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGGCATAGAC 660  
Qy 661 TTTTAACTTCAACCAATGATTTACTGAAATTAACAAATGTTGTAATTCCTGAGTGT 720  
Db 661 TTTTAACTTCAACCAATGATTTACTGAAATTAACAAATGTTGTAATTCCTGAGTGT 720  
Qy 721 TATCTACTTGTATTAAGGTAATTAATCAATCAATTAATAATCTGAGGATCAATTGCC 780  
Db 721 TATCTACTTGTATTAAGGTAATTAATCAATCAATTAATAATCTGAGGATCAATTGCC 780  
Qy 781 AGAGATTTGGGAGGGAATGTTATCAACGGTTTCATGAAATTAATTAATAAGTT 840  
Db 781 AGAGATTTGGGAGGGAATGTTATCAACGGTTTCATGAAATTAATTAATAAGTT 840  
Qy 841 ATTTCTTCAGAAAAATCAATAAGTTTGCATGTTTTTATCTTAAACATTTTAAAA 900  
Db 841 ATTTCTTCAGAAAAATCAATAAGTTTGCATGTTTTTATCTTAAACATTTTAAAA 900  
Qy 901 CCAGTGTAGATGATGTAATAGGAGCTGTGCAATTTTCTGACATATATATAAATTA 960  
Db 901 CCAGTGTAGATGATGTAATAGGAGCTGTGCAATTTTCTGACATATATATAAATTA 960  
Qy 961 TTAAGAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004  
Db 961 TTAAGAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004

RESULT 4

US-09-435-524-1

; Sequence 1, Application US/09435524  
; Patent No. 6465184

GENERAL INFORMATION:

; APPLICANT: van der Bruggen et al.

; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
THEREOF

NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/435,524  
FILING DATE: 08-No. 6465184-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/038,328  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/079,110  
FILING DATE: June 17, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: LUD 5310.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1004  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-435-524-1

Query Match 100.0%; Score 1004; DB 4; Length 1004;  
Best Local Similarity 100.0%; Pred. No. 3.7e-258;  
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCAAATTAGGCTCCGGTACTCCCGCTAGCTCTGTTCCCGCTTAGAGGACC 60  
DB 1 CGCCAAATTAGGCTCCGGTACTCCCGCTAGCTCTGTTCCCGCTTAGAGGACC 60

QY 61 AGGAGAGGGGGAGCTGGAGGCTGGAGCCTGTAACACCGTGGCTCGTCACTCTGGATG 120  
DB 61 AGGAGAGGGGGAGCTGGAGGCTGGAGCCTGTAACACCGTGGCTCGTCACTCTGGATG 120

QY 121 GTGTGGCAACAGAGATGGAGCGCAGCTGAGTGTAGAGCGCGCTGAGCGGTAGG 180  
DB 121 GTGTGGCAACAGAGATGGAGCGCAGCTGAGTGTAGAGCGCGCTGAGCGGTAGG 180

QY 181 AGTGGGGCTGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTCTGCCCAGC 240  
DB 181 AGTGGGGCTGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTCTGCCCAGC 240

QY 241 TGCTCCAGCGCAGCTGATGAGAGGAGTCCCTGTGGTGGAGCTGGAGGCTGGAGCCTG 300  
DB 241 TGCTCCAGCGCAGCTGATGAGAGGAGTCCCTGTGGTGGAGCTGGAGGCTGGAGCCTG 300

QY 301 AAGAGCGCAGAGCTGTGCTTCACTCTGAGTGTGGCAGCGGATGGAGACG 360  
DB 301 AAGAGCGCAGAGCTGTGCTTCACTCTGAGTGTGGCAGCGGATGGAGACG 360

QY 361 GCAGCTCAACAGGAGCAATAGGAGGATGAGTTCACTGTGTAGCCAGGATGGTCTC 420  
DB 361 GCAGCTCAACAGGAGCAATAGGAGGATGAGTTCACTGTGTAGCCAGGATGGTCTC 420

QY 421 GATCTCTGACCTGCTGATCGCGCCGCTTGGCCCTTCCAAAGTGGCGGAGATTACGGAT 480

DB 421 GATCTCTGACCTGCTGATCCGCCCGCTTGGCCCTTCCAAAGTGGCGGAGATTACGGAT 480  
QY 481 GTGCATTTTGTAAAGCACTTTGGAGCCACTATCAATTCGTGTGAAGAGAAATGTACCCAGA 540  
DB 481 GTGCATTTTGTAAAGCACTTTGGAGCCACTATCAATTCGTGTGAAGAGAAATGTACCCAGA 540  
QY 541 TGTATCATATCTCTGTGTGCTGAGAGCCCGCTCTTTTCCAGGATTTTCACTCATCTTCC 600  
DB 541 TGTATCATATCTCTGTGTGCTGAGAGCCCGCTCTTTTCCAGGATTTTCACTCATCTTCC 600  
QY 601 TGCTTTGTCAGAACACATTTGACCAAGCTCTCTGAAAGATGTAAGTTTACTTACGCATAGAC 660  
DB 601 TGCTTTGTCAGAACACATTTGACCAAGCTCTCTGAAAGATGTAAGTTTACTTACGCATAGAC 660  
QY 661 TTTTAAACTTCAACCAATGTATTACTGAAAAATAACAAATGTTGTAATTTCCCTGAGTGT 720  
DB 661 TTTTAAACTTCAACCAATGTATTACTGAAAAATAACAAATGTTGTAATTTCCCTGAGTGT 720  
QY 721 TATTCTACTTGTATTAAGGTAATATACATATCAATTAATAATCTGAGGATCATTTGCC 780  
DB 721 TATTCTACTTGTATTAAGGTAATATACATATCAATTAATAATCTGAGGATCATTTGCC 780  
QY 781 AGAGATTGTTGGGAGGAGAAATGTTATCAACGGTTTCATTGAAATTAATAATCCAAAGATT 840  
DB 781 AGAGATTGTTGGGAGGAGAAATGTTATCAACGGTTTCATTGAAATTAATAATCCAAAGATT 840  
QY 841 ATTTCTCTCAGAAAAATCAATTAAGTTTGGCATGTTTTTAAATCTTAAACATTTTAAAAA 900  
DB 841 ATTTCTCTCAGAAAAATCAATTAAGTTTGGCATGTTTTTAAATCTTAAACATTTTAAAAA 900  
QY 901 CCAGTGTAGATGATGTAATAGGAGCTGTGACAGTATTTCTGACATATCTATAAATA 960  
DB 901 CCAGTGTAGATGATGTAATAGGAGCTGTGACAGTATTTCTGACATATCTATAAATA 960  
QY 961 TTAAGAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004  
DB 961 TTAAGAGTCAATCAGTATTCACATCTTTTACACTAAAAAGCC 1004

## RESULT 5

US-09-382-497-1  
Sequence 1, Application US/09382497  
Patent No. 6638512  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen et al.  
TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/382,497  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: 08/389,360  
FILING DATE:  
PRIOR APPLICATION DATA: 08/079,110  
FILING DATE: June 17, 1993  
ATTORNEY/AGENT INFORMATION:

```
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5310.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1004
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-382-497-1

Query Match 100.0%; Score 1004; DB 4; Length 1004;
Best Local Similarity 100.0%; Pred. No. 3.7e-258;
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCAATTTAGGCTCCGGTATCTCCGCTGAGCTCTGTTCCTCCGGCTTAGAGGACC 60
Db 1 CGCAATTTAGGCTCCGGTATCTCCGCTGAGCTCTGTTCCTCCGGCTTAGAGGACC 60
QY 61 AGGAGAAGGGGAGCTGGAGGCTGGAAGCTGTAAACCCGTGCTCTCACTCTGGATG 120
Db 61 AGGAGAAGGGGAGCTGGAGGCTGGAAGCTGTAAACCCGTGCTCTCACTCTGGATG 120
QY 121 GTGGTGGCAACAGAGATGGCGGCGAGCTGGAGTGTAGAGGGGGGCTGAGCGGTAG 180
Db 121 GTGGTGGCAACAGAGATGGCGGCGAGCTGGAGTGTAGAGGGGGGCTGAGCGGTAG 180
QY 181 AGTGGGGCTGGAGCAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240
Db 181 AGTGGGGCTGGAGCAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240
QY 241 TGTCTCAAGCCAGCTGATGAAGAGAGTCCCTGTGTGAGCTGAGTGGAGGCTG 300
Db 241 TGTCTCAAGCCAGCTGATGAAGAGAGTCCCTGTGTGAGCTGAGTGGAGGCTG 300
QY 301 AAGACGGCAGCTCTGTGCTTCTCTGAGTTGTGCGAGCCAGCCAGCTGATGAGAG 360
Db 301 AAGACGGCAGCTCTGTGCTTCTCTGAGTTGTGCGAGCCAGCCAGCTGATGAGAG 360
QY 361 GCAGCTCAACAGAGCAATAGGAGGAGATGGAGTTTCACTGTGTACGCCAGGATG 420
Db 361 GCAGCTCAACAGAGCAATAGGAGGAGATGGAGTTTCACTGTGTACGCCAGGATG 420
QY 421 GATCTCCTGACCTGCTGATCCCGCCGCTTGGCTTCCAAAGTCCGAGATTACAGCAT 480
Db 421 GATCTCCTGACCTGCTGATCCCGCCGCTTGGCTTCCAAAGTCCGAGATTACAGCAT 480
QY 481 GTGCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCAGA 540
Db 481 GTGCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCAGA 540
QY 541 TGTATCAATATCTTGTGTGAGGAGCGGCTCTTTTCCAGGATTTCACTCACTTCC 600
Db 541 TGTATCAATATCTTGTGTGAGGAGCGGCTCTTTTCCAGGATTTCACTCACTTCC 600
QY 601 TGTCTTGTCCAGAACACATTTGACCAAGCTCTGGAAGATGTAAATTTTACGCGATAG 660
Db 601 TGTCTTGTCCAGAACACATTTGACCAAGCTCTGGAAGATGTAAATTTTACGCGATAG 660
QY 661 TTTTAACTTCAACCAATGTATTTACTGAAAAATAACAAATGTTTAAATTCCTGAGTGT 720
Db 661 TTTTAACTTCAACCAATGTATTTACTGAAAAATAACAAATGTTTAAATTCCTGAGTGT 720
QY 721 TATTTCTACTTGTATTAAGGTAAATAATACATAATCAATAATCTGAGGATCAATGCC 780
Db 721 TATTTCTACTTGTATTAAGGTAAATAATACATAATCAATAATCTGAGGATCAATGCC 780
QY 781 AGAGATTTGTGGGGGGGAAATGTTATCAACGGTTTTCAATTGAAATTAATCCAAAAGTT 840
Db 781 AGAGATTTGTGGGGGGGAAATGTTATCAACGGTTTTCAATTGAAATTAATCCAAAAGTT 840

; NAME: van der Bruggen, Pierre; Boon-Falleur,
; APPLICANT: Thierry; Coullie, Pierre; Renaud, Jean-Christophe
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: CODING FOR BAGE TUMOR REJECTION ANTIGEN PRECURSORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/079,110A
; FILING DATE: 17-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 557171man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5310
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1032 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-079-110A-1

Query Match 96.2%; Score 966; DB 1; Length 1032;
Best Local Similarity 97.3%; Pred. No. 5.1e-248;
Matches 1004; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 CGCAATTTAGGCTCTCCGGTATCTCCGCTGAGCTCTGTTCCTCCGGCTTAGAGGACC 60
Db 1 CGCAATTTAGGCTCTCCGGTATCTCCGCTGAGCTCTGTTCCTCCGGCTTAGAGGACC 60
QY 61 AGGAGAAGGGGAGCTGGAGGCTGGAAGCTGTAAACCCGTGCTCTCACTCTGGATG 120
Db 61 AGGAGAAGGGGAGCTGGAGGCTGGAAGCTGTAAACCCGTGCTCTCACTCTGGATG 120
QY 121 GTGGTGGCAACAGAGATGGCGGCGAGCTGGAGTGTAGAGGGGGGCTGAGCGGTAG 180
Db 121 GTGGTGGCAACAGAGATGGCGGCGAGCTGGAGTGTAGAGGGGGGCTGAGCGGTAG 180
QY 181 AGTGGGGCTGGAGCAAGATGGCGGCCAGAGCGGTTTTCTGGCAATGTCTGCCACG 240
```

181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTGCTGCCCGACG 240  
241 TGCTCCAAAGCCAGGCTGATGAAGAGCAGTCCCTGTGTGAGCTGGAGGTTGGAGCGTG 300  
241 TGCTCCAAAGCCAGGCTGATGAAGAGCAGTCCCTGTGTGAGCTGGAGGTTGGAGCGTG 300  
301 AAGACGGCAGAGCTGTGTGCTTCATCTTCTGAGGTTGTGGCAGCCACGGTGTAGAGACG 360  
301 AAGACGGCAGAGCTGTGTGCTTCATCTTCTGAGGTTGTGGCAGCCACGGTGTAGAGACG 360  
361 GCAGCTCAACAGAGCAATAGAGGAGAGTGGAGTTTCACTGTGTGAGCAGGATGTTCTC 420  
361 GCAGCTCAACAGAGCAATAGAGGAGAGTGGAGTTTCACTGTGTGAGCAGGATGTTCTC 420  
421 GATCTCTGACCTCGTGTGATCGCGCCGCTTGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
421 GATCTCTGACCTCGTGTGATCGCGCCGCTTGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
481 GTGCATTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
481 GTGCATTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
541 TGATCATATCTTGTGTGCTGAGGAGCGGCTTCCAGGATTTTCACTCATCTTCC 600  
541 TGATCATATCTTGTGTGCTGAGGAGCGGCTTCCAGGATTTTCACTCATCTTCC 600  
601 TGCTTTGTCAGAACACATTTGACCAAGCTTCTGAAAGATGTAAGTTTACTACGCAATAG 660  
601 TGCTTTGTCAGAACACATTTGACCAAGCTTCTGAAAGATGTAAGTTTACTACGCAATAG 660  
661 TTTTAAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAATTCCTCGATGT 720  
661 TTTTAAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAATTCCTCGATGT 720  
721 TATTTACTTTGTAATAAAGGTAATAATACATAATCATTAATAATCTGAGGATCAATGGC 780  
721 TATTTACTTTGTAATAAAGGTAATAATACATAATCATTAATAATCTGAGGATCAATGGC 780  
781 AGAGNTGTTGGGAGG-----AAATGTTATCAACG 812  
781 AGAGNTGTTGGGAGGAAATGTTATCAACGGTTTCATTAATAATGTTATCAACG 840  
813 GTTTCATTGAAATTAATCAAAAGTTATTTCTCAGAAAAATCAAAATAAAGTTTGCAT 872  
841 GTTTCATTGAAATTAATCAAAAGTTATTTCTCAGAAAAATCAAAATAAAGTTTGCAT 900  
873 GTTTCATTGAAATTAATCAAAAGTTATTTCTCAGAAAAATCAAAATAAAGTTTGCAT 932  
901 GTTTCATTGAAATTAATCAAAAGTTATTTCTCAGAAAAATCAAAATAAAGTTTGCAT 960  
933 AGTATTTCTGACATATATAATAATTAATAAAGTCAATCAGTATTCACATCTTTTA 992  
961 AGTATTTCTGACATATATAATAATTAATAAAGTCAATCAGTATTCACATCTTTTA 1020  
993 CACTAAAAAGCC 1004  
1021 CACTAAAAAGCC 1032

RESULT 7  
US-08-196-630A-1  
; Sequence 1, Application US/08196630A  
; Patent No. 5683886  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen, Pierre  
; APPLICANT: Boon-Faller, Thierry  
; TITLE OF INVENTION: ISOLATED PEPTIDES WHICH FORM  
; TITLE OF INVENTION: COMPLEXES WITH MHC MOLECULE HLA-C-CLONE 10 AND USES  
; NUMBER OF INVENTION: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue

CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/196,630A  
FILING DATE: 15-FEB-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/079,110  
FILING DATE: 17-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5683886man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5310.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1032 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-196-630A-1  
Query Match 96.2%; Score 966; DB 1; Length 1032;  
Best Local Similarity 97.3%; Pred. No. 5.1e-248;  
Matches 1004; Conservative 0; Mismatches 0; Indels 28; Gaps 1;  
QY 1 CGCCAAATTTAGGTTCTCCGGTATCTCCGGTGTAGCTGCTCTGTTCGGGCTTAGAGGACC 60  
DB 1 CGCCAAATTTAGGTTCTCCGGTATCTCCGGTGTAGCTGCTCTGTTCGGGCTTAGAGGACC 60  
QY 61 AGGAGAGGGGGAGCTGGAGCTGGAGCTGTAAACCCGTGGCTCGTCTCACTCTGGATG 120  
DB 61 AGGAGAGGGGGAGCTGGAGCTGGAGCTGTAAACCCGTGGCTCGTCTCACTCTGGATG 120  
QY 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGAGTTTAGAGGGCGGCTTAGAGGTTAGG 180  
DB 121 GTGGTGGCAACAGAGATGGCAGCGCAGCTGGAGTTTAGAGGGCGGCTTAGAGGTTAGG 180  
QY 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTGCTGCCCGACG 240  
DB 181 AGTGGGGCTGGAGCAGTAAGATGGCGCCAGAGCGGTTTTCTGGCATTTGCTGCCCGACG 240  
QY 241 TGCTCCAAAGCCAGGCTGATGAAGAGGAGTCCCTGTGTGAGCTGGAGTTGAGCGCTG 300  
DB 241 TGCTCCAAAGCCAGGCTGATGAAGAGGAGTCCCTGTGTGAGCTGGAGTTGAGCGCTG 300  
QY 301 AAGACGGCAGAGCTGTGTGCTTCATCTTCTGAGGTTGTGCGAGCCAGCGGTGTAGAGACG 360  
DB 301 AAGACGGCAGAGCTGTGTGCTTCATCTTCTGAGGTTGTGCGAGCCAGCGGTGTAGAGACG 360  
QY 361 GCAGCTCAACAGAGCAATAGAGGAGATGGAGTTTCACTGTGTGAGCAGGATGTTCTC 420  
DB 361 GCAGCTCAACAGAGCAATAGAGGAGATGGAGTTTCACTGTGTGAGCAGGATGTTCTC 420  
QY 421 GATCTCTGACCTCGTGTGATCGCGCCGCTTGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
DB 421 GATCTCTGACCTCGTGTGATCGCGCCGCTTGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
QY 481 GTGCATTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
DB 481 GTGCATTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGA 540  
QY 541 TGATCATATCTTGTGTGCTGAGGAGCGGCTTCCAGGATTTTCACTCATCTTCC 600

541	Db	TGATCATTTATCCTTTGTCTGCAGAGCGGCTCCTTTTCAGGATTTTCAGTCACATCTTC	600
601	Qy	TGCTTTGTCCAGAACACATTCGACCAAGCTCCTGAAAGATGTAAGTTTACTACGATAGAC	660
601	Db	TGCTTTGTCCAGAACACATTCGACCAAGCTCCTGAAAGATGTAAGTTTACTACGATAGAC	660
661	Qy	TTTTTAAACTTCAACCAATGATTTTACTGAAATAACAAATGTTCTGTAATTTCCCTGAGTGT	720
661	Db	TTTTTAAACTTCAACCAATGATTTTACTGAAATAACAAATGTTCTGTAATTTCCCTGAGTGT	720
721	Qy	TATCTCTACTTGTATTTAAAGGTAATAATACATAATCATTTAAAAATCTCAGGGATCATTTGCC	780
721	Db	TATCTCTACTTGTATTTAAAGGTAATAATACATAATCATTTAAAAATCTCAGGGATCATTTGCC	780
781	Qy	AGAGATGTTTGGGGAGGG-----AAATGTTTATCAACG	812
781	Db	AGAGATGTTTGGGGAGGGAAATGTTATCAACGGTTTCATTGAAATTTAAATGTTTATCAACG	840
813	Qy	GTTTTCATTGAAATTTAAATCCAAAAGTTATTTCTCTCAGAAAAATCAAAATAAAGTTTCGAT	872
841	Db	GTTTTCATTGAAATTTAAATCCAAAAGTTATTTCTCTCAGAAAAATCAAAATAAAGTTTCGAT	900
873	Qy	GTTTTTCATTCTTTAAACATTTTAAAAACCACTGTAGAATGATGTAATTAAGGACTGTGC	932
901	Db	GTTTTTCATTCTTTAAACATTTTAAAAACCACTGTAGAATGATGTAATTAAGGACTGTGC	960
933	Qy	AGTATTTCTGCACATATACTATAAATTTATTTAAAAAGTCAATCAGTATTTCAACATCTTTTAA	992
961	Db	AGTATTTCTGCACATATACTATAAATTTATTTAAAAAGTCAATCAGTATTTCAACATCTTTTAA	1020
993	Qy	CACATAAAAGCC	1004
1021	Db	CACATAAAAGCC	1032

RESULT 8  
US-08-573-186-1  
; Sequence 1, Application US/08573186  
; Patent No. 6093540  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen, Pierre; Boon-Falleur,  
; APPLICANT: Thierry; Coultre, Pierre; Renauld, Jean-Christophe  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: CODING FOR BAGE TUMOR REJECTION ANTIGEN PRECURSORS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/573,186  
; FILING DATE: 15-DEC-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/079,110  
; FILING DATE: 17-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 6093540man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5310  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 698-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:



QY 933 AGTATTTCTGACATATACTATAAAATTTATTAAGTCAATCAGTATTCACATCTTTTA 992  
 |||||||  
 Db 961 AGTATTTCTGACATATACTATAAAATTTATTAAGTCAATCAGTATTCACATCTTTTA 1020  
 |||||||  
 QY 993 CACTAAAAAGCC 1004  
 |||||||  
 Db 1021 CACTAAAAAGCC 1032  
 |||||||  
 RESULT 9  
 US-10-027-983-11  
 ; Sequence 11, Application US/10027983  
 ; Patent No. 6617162  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth W. Dobie  
 ; APPLICANT: Mark P. Roach  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION  
 ; FILE REFERENCE: RTS-0340  
 ; CURRENT APPLICATION NUMBER: US/10/027,983  
 ; CURRENT FILING DATE: 2001-12-18  
 ; NUMBER OF SEQ ID NOS: 98  
 ; SEQ ID NO 11  
 ; LENGTH: 392000  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: 137740  
 ; OTHER INFORMATION: unknown  
 ; NAME/KEY: unsure  
 ; LOCATION: 137742  
 ; OTHER INFORMATION: unknown  
 ; NAME/KEY: misc feature  
 ; LOCATION: (138122)...(138221)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: unsure  
 ; LOCATION: 145507  
 ; OTHER INFORMATION: unknown  
 ; NAME/KEY: unsure  
 ; LOCATION: 151967  
 ; OTHER INFORMATION: unknown  
 ; NAME/KEY: misc feature  
 ; LOCATION: (151967)...(1542066)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: unsure  
 ; LOCATION: 154217  
 ; OTHER INFORMATION: unknown  
 ; NAME/KEY: misc feature  
 ; LOCATION: (164037)...(164136)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: misc feature  
 ; LOCATION: (174657)...(174756)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: misc feature  
 ; LOCATION: (186224)...(186323)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: misc feature  
 ; LOCATION: (195242)...(195341)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: unsure  
 ; LOCATION: 202703  
 ; OTHER INFORMATION: unknown  
 ; NAME/KEY: misc feature  
 ; LOCATION: (202771)...(202870)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: misc feature  
 ; LOCATION: (206246)...(215602)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: misc feature  
 ; LOCATION: (218126)...(218225)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; NAME/KEY: misc feature

LOCATION: (220360)...(220459)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (222717)...(222816)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (223981)...(224080)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (227487)...(227586)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (230157)...(230256)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (232299)...(232398)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (236552)...(236651)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: misc feature  
 LOCATION: (238789)...(248788)  
 OTHER INFORMATION: n = A,T,C or G  
 NAME/KEY: exon  
 LOCATION: (118288)...(119101)  
 OTHER INFORMATION: exon 1C  
 NAME/KEY: exon:intron junction  
 LOCATION: (151129)...(151130)  
 OTHER INFORMATION: exon 5:intron 5  
 NAME/KEY: exon:intron junction  
 LOCATION: (299248)...(299249)  
 OTHER INFORMATION: exon 9:intron 9  
 NAME/KEY: exon:intron junction  
 LOCATION: (348578)...(348579)  
 OTHER INFORMATION: exon 10:intron 10  
 NAME/KEY: intron  
 LOCATION: (348579)...(381838)  
 OTHER INFORMATION: intron 10  
 NAME/KEY: intron:exon junction  
 LOCATION: (386185)...(386186)  
 OTHER INFORMATION: intron 11:exon 12  
 US-10-027-983-11  
 Query Match 8.3%; Score 83; DB 4; Length 392000;  
 Best Local Similarity 89.9%; Pred. No. 2.4e-11;  
 Matches 89; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 386 AGATGGAGTTTCACTGTGTGAGCCAGGATGCTCGATCTCTGACCTCTGATCCGCC 445  
 |||||||  
 Db 252283 AGACGGAGTTTCACTGTGTGAGCCAGGATGCTCTGATCTCTGACCTCTGATCCACC 252342  
 |||||||  
 QY 446 GCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTGC 484  
 |||||||  
 Db 252343 GCCTAGCCTTCCAAAGTCTGGGATTACAGGATGAGC 252381  
 |||||||  
 RESULT 10  
 US-09-535-008-58  
 ; Sequence 58, Application US/09535008  
 ; Patent No. 6465829  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wong, Alexander K.C.  
 ; APPLICANT: Tavtigian, Sean V.  
 ; APPLICANT: Teng, David H.-F.  
 ; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
 ; FILE REFERENCE: 2318-259  
 ; CURRENT APPLICATION NUMBER: US/09/535,008  
 ; CURRENT FILING DATE: 2000-03-23  
 ; EARLIER APPLICATION NUMBER: U.S. 60/125,806  
 ; EARLIER FILING DATE: 1999-03-23  
 ; NUMBER OF SEQ ID NOS: 77  
 ; SOFTWARE: PatentIn Ver. 2.0



```
; SEQ ID NO 58
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-535-008-58

Query Match      8.2%; Score 82.4; DB 4; Length 859;
Best Local Similarity 89.0%; Pred. No. 2e-12;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 385 GAGATGGAGTTTCACTGTGTACGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCC 444
    |||||
Db 80 GAGACGGGGTTTCACTGTGTGTAGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCC 139
    |||||

QY 445 CGCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTC 484
    |||||
Db 140 CGCCTTGGCCTTCCAAAGTCTGGGATTACGGCGGTGAGC 179
    |||||

RESULT 11
US-09-439-261-34
; Sequence 34, Application US/09439261
; Patent No. 6428990
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pardeep
; APPLICANT: Leonard, Amanda E.
; APPLICANT: Huang, Yung-Sheng
; TITLE OF INVENTION: HUMAN DESATURASE GENE AND USES THEREOF
; FILE REFERENCE: 6295.US.P2
; CURRENT FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 08/833,610
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: PCT/US98/07422
; PRIOR FILING DATE: 1998-04-10
; PRIOR APPLICATION NUMBER: US 09/227,613
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 4698
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-261-34

Query Match      8.2%; Score 82.4; DB 4; Length 4698;
Best Local Similarity 89.0%; Pred. No. 4.4e-12;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 385 GAGATGGAGTTTCACTGTGTACGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCC 444
    |||||
Db 3324 GAGACGGGGTTTCACTGTGTGTAGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCACC 3383
    |||||

QY 445 CGCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTC 484
    |||||
Db 3384 CGCCTTGGCCTTCCAAAGTCTGGGATTACAGCGGTGAGC 3423
    |||||

RESULT 12
US-09-227-613-33
; Sequence 33, Application US/09227613A
; Patent No. 6432684
; GENERAL INFORMATION:
; APPLICANT: MUKERJI, Pradip
; APPLICANT: LEONARD, Amanda E.
; APPLICANT: HUANG, Yung-Sheng
; TITLE OF INVENTION: HUMAN DESATURASE GENE AND USES THEREOF
; FILE REFERENCE: 6295.US.P1
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 08/833,610
; PRIOR FILING DATE: 1997-04-11
```

```
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 4698
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-227-613-33

Query Match      8.2%; Score 82.4; DB 4; Length 4698;
Best Local Similarity 89.0%; Pred. No. 4.4e-12;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 385 GAGATGGAGTTTCACTGTGTACGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCC 444
    |||||
Db 3324 GAGACGGGGTTTCACTGTGTGTAGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCACC 3383
    |||||

QY 445 CGCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTC 484
    |||||
Db 3384 CGCCTTGGCCTTCCAAAGTCTGGGATTACAGCGGTGAGC 3423
    |||||

RESULT 13
US-09-844-634-17/c
; Sequence 17, Application US/09844634
; Patent No. 6410324
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF TUMOR NECROSIS FACTOR RECEPTOR 2 EXPRESSION
; FILE REFERENCE: RTS-0216
; CURRENT APPLICATION NUMBER: US/09/844,634
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 174
; SEQ ID NO 17
; LENGTH: 15602
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-844-634-17

Query Match      8.2%; Score 82.4; DB 4; Length 15602;
Best Local Similarity 89.0%; Pred. No. 7.7e-12;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 385 GAGATGGAGTTTCACTGTGTACGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCC 444
    |||||
Db 1945 GAGATGGGGTTTCACTGTGTGTAGCCAGGATGGTCTCGATCTCTGACCTCGTGATCCGCC 1886
    |||||

QY 445 CGCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTC 484
    |||||
Db 1885 CCCCTCGGCCTTCCAAAGTCTGGGATTACAGCGGTGAGC 1846
    |||||

RESULT 14
US-08-306-691B-19
; Sequence 19, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
```

```
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-306-691B-19

Query Match      8.2%; Score 82.4; DB 1; Length 35100;
Best Local Similarity 89.0%; Pred. No. 1.1e-11;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 385 GAGATGGAGTTTCACTGTGTGTCAGCCAGGATGGTCTCGATCTCTGACCTCGTGTATCGGCC 444
Db 28167 GAGATGGGGTTTACCAATGTTAGCCAGGATGGTCTCGATCTCTGACCTCGTGTATCGGCC 444
QY 445 CGCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTGC 484
Db 28227 CACCTTGGCCTCCCAAGTCTGAGATTACAGGCAATGAGC 28266

Search completed: July 9, 2004, 10:09:37
Job time : 110 secs
```

```
Query Match      8.2%; Score 82.4; DB 1; Length 35100;
Best Local Similarity 89.0%; Pred. No. 1.1e-11;
Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 385 GAGATGGAGTTTCACTGTGTGTCAGCCAGGATGGTCTCGATCTCTGACCTCGTGTATCGGCC 444
Db 28167 GAGATGGGGTTTACCAATGTTAGCCAGGATGGTCTCGATCTCTGACCTCGTGTATCGGCC 28226
QY 445 CGCCTTGGCCTTCCAAAGTCCGAGATTACAGCGATGTGC 484
Db 28227 CACCTTGGCCTCCCAAGTCTGAGATTACAGGCAATGAGC 28266

RESULT 15
PCT-US93-06251-19
; Sequence 19, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 08:57:39 ; Search time 694 Seconds  
(without alignments)  
7050.978 Million cell updates/sec

Title: US-10-081-108-1  
Perfect score: 1004  
Sequence: 1 CGCCAAATTAGGCTCCGG.....ATCTTTACACTAAAAGCC 1004

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3183909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters: 6367818

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
12: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
13: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
14: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:  
15: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:  
16: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:  
17: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:  
18: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1004	100.0	1004	15	US-10-207-655-80
2	1004	100.0	1004	15	US-10-081-108-1
3	524.4	52.2	1162	17	US-10-363-408-1
4	524.4	52.2	13928	13	US-10-362-916-1
5	330.4	32.9	524	15	US-10-029-386-1510
6	330.4	32.9	524	15	US-10-029-386-8714
7	164.6	16.4	345	15	US-10-029-386-18339
8	164.6	16.4	538	15	US-10-029-386-4583
9	164.2	16.4	172	15	US-10-029-386-22449
10	149.8	14.9	153	15	US-10-029-386-15211
11	149.8	14.9	6292	13	US-10-221-714A-461
12	93.6	9.3	31516	13	US-10-087-192-1252
13	88.4	8.8	407	13	US-10-027-632-128026
14	88.4	8.8	407	13	US-10-027-632-128027

C 15	88.4	8.8	407	13	US-10-027-632-128028	Sequence 128028,
C 16	88.4	8.8	407	13	US-10-027-632-131415	Sequence 131415,
C 17	88.4	8.8	407	16	US-10-027-632-128026	Sequence 128026,
C 18	88.4	8.8	407	16	US-10-027-632-128027	Sequence 128027,
C 19	88.4	8.8	407	16	US-10-027-632-128028	Sequence 128028,
C 20	88.4	8.8	407	16	US-10-027-632-131415	Sequence 131415,
C 21	87	8.7	130	9	US-09-764-877-2836	Sequence 2836, Ap
C 22	87	8.7	130	9	US-09-764-877-2835	Sequence 2835, Ap
C 23	87	8.7	130	9	US-09-764-877-2835	Sequence 2835, Ap
C 24	87	8.7	130	9	US-09-764-877-2835	Sequence 2835, Ap
C 25	86.6	8.6	476	9	US-09-998-598-1168	Sequence 1168, Ap
C 26	86.2	8.6	72332	12	US-10-052-482-58	Sequence 58, Appl
C 27	86	8.6	6292	13	US-10-221-714A-462	Sequence 462, App
C 28	85.6	8.5	411	9	US-09-954-456-1450	Sequence 1450, Ap
C 29	85.6	8.5	411	9	US-09-880-107-445	Sequence 445, Appl
C 30	85.6	8.5	113000	15	US-10-376-566-16	Sequence 16, Appl
C 31	85.2	8.5	91760	13	US-10-087-192-844	Sequence 844, App
C 32	84.6	8.4	2187	16	US-10-108-260A-1171	Sequence 1171, Ap
C 33	84.4	8.4	667	10	US-09-764-891-6601	Sequence 6601, Ap
C 34	84.4	8.4	667	10	US-09-764-891-6603	Sequence 6603, Ap
C 35	84.4	8.4	667	15	US-10-091-572-465	Sequence 465, App
C 36	84.4	8.4	667	15	US-10-091-572-467	Sequence 467, App
C 37	84.4	8.4	1870	13	US-10-027-632-97588	Sequence 97588, A
C 38	84.4	8.4	1870	16	US-10-027-632-97588	Sequence 97588, A
C 39	84.4	8.4	312477	17	US-10-317-883A-12	Sequence 12, Appl
C 40	84.2	8.4	125042	13	US-10-087-192-1240	Sequence 1240, Ap
C 41	84	8.4	358	10	US-09-918-995-29818	Sequence 29818, A
C 42	84	8.4	846	13	US-10-027-632-157125	Sequence 157125,
C 43	84	8.4	846	13	US-10-027-632-157126	Sequence 157126,
C 44	84	8.4	846	16	US-10-027-632-157125	Sequence 157125,
C 45	84	8.4	846	16	US-10-027-632-157126	Sequence 157126,

ALIGNMENTS

RESULT 1  
US-10-207-655-80  
; Sequence 80, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069, 401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 80  
; LENGTH: 1004  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-207-655-80

Query Match	100.0%	Score 1004;	DB 15;	Length 1004;
Best Local Similarity	100.0%	Pred. No. 2.8e-247;		
Matches 1004;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	CGCCAAATTAGGCTCCGGTATCTCCGCTGAGCTGCTCTGTTCCCGGCTTAGAGACC	60	
Db	1	CGCCAAATTAGGCTCCGGTATCTCCGCTGAGCTGCTCTGTTCCCGGCTTAGAGACC	60	
Qy	61	AGGAGAGCGGGAGCTGGAGGCTGAGGCTGTAACCGTGGCTGCTCACTCTGGATG	120	
Db	61	AGGAGAGCGGGAGCTGGAGGCTGAGGCTGTAACCGTGGCTGCTCACTCTGGATG	120	
Qy	121	GTGTGGCAACAGAGATGGCAGCGAGCTGAGTGTGTAGGAGCGCGCTGAGCGTAGG	180	
Db	121	GTGTGGCAACAGAGATGGCAGCGAGCTGAGTGTGTAGGAGCGCGCTGAGCGTAGG	180	
Qy	181	AGTGGGCTGAGCAGTAAGTGGCGCCAGAGCGGTTTCTGGCATTGCTGCCAGC	240	

Db 181 AGTGGGGCTGGAGCAGTAAGATGGCGGCGAGAGCGGTTTTCTGGCATTTGTCTGCCAGC 240  
Qy 241 TGCTTCAAGCCAGGCTGATGAAGAGAGAGTCCCTGTGTGTGAGCTGGAGGTTGGAGCCTG 300  
Db 241 TGCTTCAAGCCAGGCTGATGAAGAGAGAGTCCCTGTGTGTGAGCTGGAGGTTGGAGCCTG 300  
Qy 301 AAGAGCGCACAGCTCTGTGCTTCTATCTTCTGAGGTTGTGGCAGCCAGCGTGATGGAGCG 360  
Db 301 AAGAGCGCACAGCTCTGTGCTTCTATCTTCTGAGGTTGTGGCAGCCAGCGTGATGGAGCG 360  
Qy 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGAGGTTTCACTGTGTGAGCAGGATGTCTC 420  
Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGAGGTTTCACTGTGTGAGCAGGATGTCTC 420  
Qy 421 GATCTCTGACCTGTGATCCGCGGCTTGGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
Db 421 GATCTCTGACCTGTGATCCGCGGCTTGGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
Qy 481 GTGCATTTTGTAGACCTTTGGAGGCACTATCAAAATGCTGTGAAGAGAAATGTACCAGA 540  
Db 481 GTGCATTTTGTAGACCTTTGGAGGCACTATCAAAATGCTGTGAAGAGAAATGTACCAGA 540  
Qy 541 TGATCATTTATCTTGTGCTGAGGAGCGGCTTCCAGGATTTTCACTGTGTGAGCAGGATGTCTC 600  
Db 541 TGATCATTTATCTTGTGCTGAGGAGCGGCTTCCAGGATTTTCACTGTGTGAGCAGGATGTCTC 600  
Qy 601 TGCTTTGTCAGACACATTTGACCGAGCTCTGAAAGATGTAAGTTTACTACCATAGAC 660  
Db 601 TGCTTTGTCAGACACATTTGACCGAGCTCTGAAAGATGTAAGTTTACTACCATAGAC 660  
Qy 661 TTTTAAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAAATTCCTGTAGTGT 720  
Db 661 TTTTAAACTTCAACCAATGATTTTACTGAAATTAACAAATGTTGTAAATTCCTGTAGTGT 720  
Qy 721 TATTCATCTGTATTAAGAGTAAATATACATTAATTAATTAATTAATTAATTAATTAATTA 780  
Db 721 TATTCATCTGTATTAAGAGTAAATATACATTAATTAATTAATTAATTAATTAATTAATTA 780  
Qy 781 AGAGATTGTTGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAATTAATTAATTAATTA 840  
Db 781 AGAGATTGTTGGGAGGGAATGTTATCAACGGTTTCATTGAAATTAATTAATTAATTAATTA 840  
Qy 841 ATTTCCCTCAGAAAAATCAATTAAGTTTGCATGTTTTTATTTCTTAAACATTTTAAAAA 900  
Db 841 ATTTCCCTCAGAAAAATCAATTAAGTTTGCATGTTTTTATTTCTTAAACATTTTAAAAA 900  
Qy 901 CCAGTGTAGATGATGATTAATAGGAGCTGTGAGTATTTCTGACATATATATAAATTA 960  
Db 901 CCAGTGTAGATGATGATTAATAGGAGCTGTGAGTATTTCTGACATATATATAAATTA 960  
Qy 961 TTAAGAGTCAATCAGTATTTCAACATCTTTTACATAAAGCC 1004  
Db 961 TTAAGAGTCAATCAGTATTTCAACATCTTTTACATAAAGCC 1004

## RESULT 2

US-10-081-108-1

; Sequence 1, Application US/10081108

; Publication No. US20030138854A1

; GENERAL INFORMATION:

; APPLICANT: van der Bruggen et al.

; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

; WITH MHC MOLECULE HLA-Cw\*1601 AND USES

; THEREOF

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe &amp; Lynch

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/081,108  
; FILING DATE: 20-Feb-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/435,524  
; FILING DATE: 08-No. US20030138854A1-1999  
; APPLICATION NUMBER: 09/038,328  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/079,110  
; FILING DATE: June 17, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pasqualini, Patricia A.  
; REGISTRATION NUMBER: 34,894  
; REFERENCE/DOCKET NUMBER: LUD 5310.2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1004  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-081-108-1

Query Match 100.0%; Score 1004; DB 15; Length 1004;

Best Local Similarity 100.0%; Pred. No. 2.8e-247;

Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCAAATTTAGGTTCTCGGTTATCTCCGCTGAGCTCTGTTCCTCGGCTTAGAGACC 60  
Db 1 CGCCAAATTTAGGTTCTCGGTTATCTCCGCTGAGCTCTGTTCCTCGGCTTAGAGACC 60  
Qy 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACCCGTGGTCTCACTCTGGATG 120  
Db 61 AGGAGAGGGGAGCTGGAGCTGGAGCTGTAAACCCGTGGTCTCACTCTGGATG 120  
Qy 121 GTGTGGCAACAGAGATGGCAGCGAGCTGTGAGTTAGAGGGGCGCTGAGCGGTAGG 180  
Db 121 GTGTGGCAACAGAGATGGCAGCGAGCTGTGAGTTAGAGGGGCGCTGAGCGGTAGG 180  
Qy 181 AGTGGGGCTGGAGCAGTAAGATGCGGCGCAGAGCGGTTTTTCTGGCATTTGTCTCCCGCAGC 240  
Db 181 AGTGGGGCTGGAGCAGTAAGATGCGGCGCAGAGCGGTTTTTCTGGCATTTGTCTCCCGCAGC 240  
Qy 241 TGCTTCAAGCCAGGCTGATGAAGAGGAGTCCCTGTGTGAGCTGGAGGTTGGAGCCTG 300  
Db 241 TGCTTCAAGCCAGGCTGATGAAGAGGAGTCCCTGTGTGAGCTGGAGGTTGGAGCCTG 300  
Qy 301 AAGAGCGCACAGCTGTGCTTCTATCTTCTGAGGTTGGCGAGCCAGGTTGATGGAGACG 360  
Db 301 AAGAGCGCACAGCTGTGCTTCTATCTTCTGAGGTTGGCGAGCCAGGTTGATGGAGACG 360  
Qy 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTCAAGGATGTGTCTC 420  
Db 361 GCAGCTCAACAGGAGCAATAGGAGGAGATGGAGTTTCACTGTGTCAAGGATGTGTCTC 420  
Qy 421 GATCTCTGACCTGTGATCCGCGGCTTGGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
Db 421 GATCTCTGACCTGTGATCCGCGGCTTGGGCTTCCAAAGTGGCGAGATTACAGCGAT 480  
Qy 481 GTGCATTTTGTAAAGCATTGTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCAGA 540  
Db 481 GTGCATTTTGTAAAGCATTGTGGAGCCACTATCAAAATGCTGTGAAGAGAAATGTACCAGA 540  
Qy 541 TGATCATTTATCTTGTGCTGAGGAGCGGCTTCCAGGATTTTCACTGTGTGAGCAGGATGTCTC 600

Db 541 TGTATCATTTATCTTGTCTGCAGGAGCGGCTCTTTTCAAGGATTTTCAGTCACATCTTC 600  
QY 601 TGCTTTCTCAGAACACATTTGACCAAGCTCTCTGAAGAGATGTAAGTTTACTACGCATAGAC 660  
Db 601 TGCTTTCTCAGAACACATTTGACCAAGCTCTCTGAAGAGATGTAAGTTTACTACGCATAGAC 660  
QY 661 TTTTAAACTTCAACCAATGATTTACTGAAAAAATAACAAATGTTTGTAAATTCCTGAGTGT 720  
Db 661 TTTTAAACTTCAACCAATGATTTACTGAAAAAATAACAAATGTTTGTAAATTCCTGAGTGT 720  
QY 721 TATTTCTACTTGTATTAAGAGTAAATATACATATCATTAATCAATTAATCAATTAATCAATTAAT 780  
Db 721 TATTTCTACTTGTATTAAGAGTAAATATACATATCATTAATCAATTAATCAATTAATCAATTAAT 780  
QY 781 AGAGATGTTGGGAGGAGAAATGTTTCAACGGTTCATTAAGATTAATTAATTAATTAATTAAT 840  
Db 781 AGAGATGTTGGGAGGAGAAATGTTTCAACGGTTCATTAAGATTAATTAATTAATTAATTAAT 840  
QY 841 ATTTCTCAGAAAAATCAATAAAGTTTCAACGGTTCATTAAGATTAATTAATTAATTAATTAAT 900  
Db 841 ATTTCTCAGAAAAATCAATAAAGTTTCAACGGTTCATTAAGATTAATTAATTAATTAATTAAT 900  
QY 901 CCACTGTAGATGATTAATAGGAGCTGTGCAAGTATTTCTGACATATATTAATTAATTAAT 960  
Db 901 CCACTGTAGATGATTAATAGGAGCTGTGCAAGTATTTCTGACATATATTAATTAATTAAT 960  
QY 961 TTAAGAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004  
Db 961 TTAAGAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004

RESULT 3  
US-10-363-408-1/c  
; Sequence 1, Application US/10363408  
; Publication No. US2004007042A1  
; GENERAL INFORMATION:  
; APPLICANT: XIA, Jiahui  
; TITLE OF INVENTION: A Cell Line Expressing Mutated Human Tissue-Type Plasminogen  
; TITLE OF INVENTION: Activator, Its Construction Strategy and Method of Expressed  
; TITLE OF INVENTION: Protein Preparation  
; FILE REFERENCE: 4596-0103P  
; CURRENT APPLICATION NUMBER: US/10/363,408  
; CURRENT FILING DATE: 2003-10-27  
; PRIOR APPLICATION NUMBER: PCT/CN01/00127  
; PRIOR FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: PCT/CN00/00260  
; PRIOR FILING DATE: 2000-09-04  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 11162  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(11162)  
; OTHER INFORMATION: target gene vector  
US-10-363-408-1

Query Match 52.2%; Score 524.4; DB 17; Length 11162;  
Best Local Similarity 98.9%; Pred. No. 2.2e-123;  
Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 471 TTACAGCGATGTCATTTTGTAAAGCCTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAA 530  
Db 6578 TTTTACGCAATGTCATTTTGTAAAGCCTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAA 6519  
QY 531 TGATCCAGATGATCATTTATCTTGTGTGAGGAGCGGCTCTTTTCAAGATTTTCAAGT 590  
Db 6518 TGATCCAGATGATCATTTATCTTGTGTGAGGAGCGGCTCTTTTCAAGATTTTCAAGT 6459  
QY 591 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACT 650

Db 6458 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACT 6399  
QY 651 AGCGATAGACTTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTGTAATTT 710  
Db 6398 AGCGATAGACTTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTGTAATTT 6339  
QY 711 CCCTGAGTGTATTCTACTTCTTATTAAGAGTAAATATACATAATCATTAATAATCTGAGG 770  
Db 6338 CCCTGAGTGTATTCTACTTCTTATTAAGAGTAAATATACATAATCATTAATAATCTGAGG 6279  
QY 771 GATCATTTCCAGAGATTTGTTGGGAGGGAATGTTTATCAACGGTTCATTTGAAATTAAT 830  
Db 6278 GATCATTTCCAGAGATTTGTTGGGAGGGAATGTTTATCAACGGTTCATTTGAAATTAAT 6219  
QY 831 CCAAAAAAGTATTCTCTCAGAAAAATCAATAAAGTTTGAAGTTTGTATTTTCTTAAAC 890  
Db 6218 CCAAAAAAGTATTCTCTCAGAAAAATCAATAAAGTTTGTATTTTATCTTAAAC 6159  
QY 891 ATTTTAAAAACACACTGTAGAATGATGTAATAGGAGCTGTGACATATTTTCTGACATATAC 950  
Db 6158 ATTTTAAAAACACACTGTAGAATGATGTAATAGGAGCTGTGCAATATTTCTGACATATAC 6099  
QY 951 TATAAAATTTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 1004  
Db 6098 TATAAAATTTTAAAAAGTCAATCAGTATTTCAACATCTTTTACACTAAAAAGCC 6045

RESULT 4  
US-10-362-916-1/c  
; Sequence 1, Application US/10362916  
; Publication No. US20040038919A1  
; GENERAL INFORMATION:  
; APPLICANT: XIA, Jiahui  
; TITLE OF INVENTION: A No. US20040038919A1 Gene Therapy Agent For Hemophilia B And  
; TITLE OF INVENTION: Method  
; FILE REFERENCE: 4596-0102P  
; CURRENT APPLICATION NUMBER: US/10/362,916  
; CURRENT FILING DATE: 2003-06-30  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 13928  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-362-916-1

Query Match 52.2%; Score 524.4; DB 13; Length 13928;  
Best Local Similarity 98.9%; Pred. No. 2.6e-123;  
Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 471 TTACAGCGATGTCATTTTGTAAAGCCTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAA 530  
Db 9344 TTTTACGCAATGTCATTTTGTAAAGCCTTTGGAGCCACTATCAAAATGCTGTGAAGAGAAA 9285  
QY 531 TGATCCAGATGATCATTTATCTTGTGTGAGGAGCGGCTCTTTTCAAGATTTTCAAGT 590  
Db 9284 TGATCCAGATGATCATTTATCTTGTGTGAGGAGCGGCTCTTTTCAAGATTTTCAAGT 9225  
QY 591 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACT 650  
Db 9224 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTCTGAAAGATGTAAGTTTACT 9165  
QY 651 AGCGATAGACTTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTGTAATTT 710  
Db 9164 AGCGATAGACTTTTAAACTTCAACCAATGTAATTTACTGAAAAATAACAAATGTTGTAATTT 9105  
QY 711 CCCTGAGTGTATTCTACTTGTATTAAGAGTAAATATACATAATCATTAATAATCTGAGG 770  
Db 9104 CCCTGAGTGTATTCTACTTGTATTAAGAGTAAATATACATAATCATTAATAATCTGAGG 9045  
QY 771 GATCATTTCCAGAGATTTGTTGGGAGGGAATGTTTCAACGGTTCATTTGAAATTAAT 830  
Db 9044 GATCATTTCCAGAGATTTGTTGGGAGGGAATGTTTCAACGGTTCATTTGAAATTAAT 8985

QY 831 CCAAAAAGTTATTCCTCAGAAAAATCAATAAAGTTTGCATGTTTTTTTATTTCTTAAAC 890  
Db 8984 CCAAAAAGTTATTCCTCAGAAAAATCAATAAAGTTTTCATGTTTTTTTATTTCTTAAAC 8925  
QY 891 ATTTTAAAAACCACTGTAGAAATCATGTAAATAGGAGCTGCAATGTTTCTGCATATAC 950  
Db 8924 ATTTTAAAAACCACTGTAGAAATCATGTAAATAGGAGCTGCAATGTTTCTGCATATAC 8865  
QY 951 TATAAATTTATTAAGAGTCAATCAGTATTCACATCTTTTACACATAAAGCC 1004  
Db 8864 TATAAATTTATTAAGAGTCAATCAGTATTCACATCTTTTACACATAAAGCC 8811

## RESULT 5

US-10-029-386-1510/c  
; Sequence 1510, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 1510  
; LENGTH: 524  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR21 2.0  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9  
; OTHER INFORMATION: SWISSPROT HIT: P55200, EVALUE 9.00e-04  
; OTHER INFORMATION: NT HIT: AL163201.2, EVALUE 0.00e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: A1138404.1, EVALUE 0.00e+00  
US-10-029-386-1510

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
Best Local Similarity 98.2%; Pred. No. 2.2e-74;  
Matches 334; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 471 TTACAGCGATGTGCAATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
Db 340 TTTACAGCGATGTGCAATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 281  
QY 531 TGTACCCAGATGTATCATATTCCTTGTGCTGCAGAGCCGCGCTCTTTTCAGGATTTTCA 590  
Db 280 TGTACCCAGATGTATCATATTCCTTGTGCTGCAGAGCCGCGCTCTTTTCAGGATTTTCA 221  
QY 591 CACATCTTCTGCTTTGTGTCAGAGCAACATTGACCAAGCTCTGAAAGATGTAAATTTTACT 650  
Db 220 CACATCTTCTGCTTTGTGTCAGAGCAACATTGACCAAGCTCTGAAAGATGTAAATTTTACT 161  
QY 651 ACGCATAGACTTTTAAACTTCAACCAATGTATTTACTGAAAAATAACAAATGTTGTAATTT 710  
Db 160 ACGCATAGACTTTTAAACTTCAACCAATGTATTTACTGAAAAATAACAAATGTTGTAATTT 101  
QY 711 CCCTGAGTGTATTTCTACTGTATTAAGGTAAATAATACATATCAATTAATAATCTGAGG 770  
Db 100 CCCTGAGTGTATTTCTACTGTATTAAGGTAAATAATACATATCAATTTTAAATCTGAGG 41  
QY 771 GATCATGCCAGAGATTTGGGAGGAGAAATGTTATCAA 810  
Db 40 GATCATGCCAGAGATTTGGGAGGAGAAATGTTATCAA 1

## RESULT 6

US-10-029-386-8714  
; Sequence 8714, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 8714  
; LENGTH: 524  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AF254983.1  
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 0.95  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95  
; OTHER INFORMATION: EST\_HUMAN HIT: A1138404.1, EVALUE 0.00e+00  
; OTHER INFORMATION: NT HIT: AL163201.2, EVALUE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: P55200, EVALUE 9.00e-04  
US-10-029-386-8714

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
Best Local Similarity 98.2%; Pred. No. 2.2e-74;  
Matches 334; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 471 TTACAGCGATGTGCAATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
Db 185 TTTACAGCGATGTGCAATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 244  
QY 531 TGTACCCAGATGTATCATATTCCTTGTGCTGCAGAGCCGCGCTCTTTTCAGGATTTTCA 590  
Db 245 TGTACCCAGATGTATCATATTCCTTGTGCTGCAGAGCCGCGCTCTTTTCAGGATTTTCA 304  
QY 591 CACATCTTCTGCTTTGTGTCAGAGCAACATTGACCAAGCTCTGAAAGATGTAAATTTTACT 650  
Db 305 CACATCTTCTGCTTTGTGTCAGAGCAACATTGACCAAGCTCTGAAAGATGTAAATTTTACT 364  
QY 651 ACGCATAGACTTTTAAACTTCAACCAATGTATTTACTGAAAAATAACAAATGTTGTAATTT 710  
Db 365 ACGCATAGACTTTTAAACTTCAACCAATGTATTTACTGAAAAATAACAAATGTTGTAATTT 424  
QY 711 CCCTGAGTGTATTTCTACTGTATTAAGGTAAATAATACATATCAATTTTAAATCTGAGG 770  
Db 425 CCCTGAGTGTATTTCTACTGTATTAAGGTAAATAATACATATTTTAAATCTGAGG 484  
QY 771 GATCATGCCAGAGATTTGGGAGGAGAAATGTTATCAA 810  
Db 485 GATCATGCCAGAGATTTGGGAGGAGAAATGTTATCAA 524

## RESULT 7

US-10-029-386-18339/c  
; Sequence 18339, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16339
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049849.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: NT HIT: g116165974, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE890756.1, EVALUE 7.00e-65
; OTHER INFORMATION: SWISSPROT HIT: Q13072, EVALUE 2.00e-06
US-10-029-386-18339
```

```
Query Match 16.4%; Score 164.6; DB 15; Length 345;
Best Local Similarity 97.7%; Pred. No. 6.7e-32;
Matches 167; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 215 GGTGTTCTGGCATTGTCTCCAGCTGCTCCAGCCAGCTGATGAAGAGGATCCCC 274
DB 282 GGTGTTCTGGCATTGTCTCCAGCTGCTCCAGCCAGCTGATGAAGAGGATCCCC 223
QY 275 TGTGTCAGCTGGAGGTTGAGCTGAAGACGCGACAGCTCTGTCTTCTTCTGAGG 334
DB 222 TGTGTCAGCTGGAGGTTGAGCTGAAGATGCGACAGCTCTGTGATTCTTCTGCGG 163
QY 335 TTGTGGCAGCCAGGTTGATGAGCAGCGCAGCTCAACAGGAGCAATAGGAGG 385
DB 162 TTGTGGCAGCCAGGTTGATGAGCAGCGCAGCTCAACAGGAGCAATAGGAGG 112
```

## RESULT 8

```
US-10-029-386-4583/c
; Sequence 4583, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4583
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL049849.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: Q13072, EVALUE 4.00e-06
; OTHER INFORMATION: NT HIT: g114763849, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE890756.1, EVALUE 1.00e-64
; OTHER INFORMATION: EST_HUMAN HIT: BE890756.1, EVALUE 1.00e-64
US-10-029-386-4583
```

```
Query Match 16.4%; Score 164.6; DB 15; Length 538;
Best Local Similarity 97.7%; Pred. No. 9e-32;
Matches 167; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 215 GGTGTTCTGGCATTGTCTCCAGCTGCTCCAGCCAGCTGATGAAGAGGATCCCC 274
```

```
DB 366 GGTGTTCTGGCATTGTCTCCAGCTGCTCCAGCCAGCTGATGAAGAGGAGTCCCC 307
QY 275 TGTGTCAGCTGGAGGTTGAGCTGAAGACGCGCAGCTCTGTCTTCTTCTGAGG 334
DB 306 TGTGTCAGCTGGAGGTTGAGCTGAAGATGCGACAGCTCTGTGATTCTTCTGCGG 247
QY 335 TTGTGGCAGCCAGGTTGATGAGCAGCGCAGCTCAACAGGAGCAATAGGAGG 385
DB 246 TTGTGGCAGCCAGGTTGATGAGCAGCGCAGCTCAACAGGAGCAATAGGAGG 196
```

## RESULT 9

```
US-10-029-386-22449
; Sequence 22449, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22449
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF254983.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95
; OTHER INFORMATION: SWISSPROT HIT: P55200, EVALUE 2.00e-04
; OTHER INFORMATION: EST_HUMAN HIT: A1138404.1, EVALUE 6.00e-82
; OTHER INFORMATION: NT HIT: A1163201.2, EVALUE 3.00e-92
US-10-029-386-22449
```

```
Query Match 16.4%; Score 164.2; DB 15; Length 172;
Best Local Similarity 98.2%; Pred. No. 5.5e-32;
Matches 166; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 471 TTACAGCAGTGTCATTTTGTAAAGCATTTCGAGCCACTATCAATGCTGTGAAGAGAA 530
DB 4 TTTACGAGTGTCATTTTGTAAAGCATTTCGAGCCACTATCAATGCTGTGAAGAGAA 63
QY 531 TGTACCCAGATGATCATTTCTTGTCTGTCAGGAGCCGGCTCTTTTCAGGATTTCAAT 590
DB 64 TGTACCCAGATGATCATTTCTTGTCTGTCAGGAGCCGACACTTTTCAGGATTTCAAT 123
QY 591 CACATCTTCTGCTTTGTCCGAGAACACATTTCACCAAGCTCTTGAAGAT 639
DB 124 CACATCTTCTGCTTTGTCCGAGAACACATTTCACCAAGCTCTTGAAGAT 172
```

## RESULT 10

```
US-10-029-386-15211/c
; Sequence 15211, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
```



```
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15211
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21.2.0
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: NT HIT: 9114756516, EVALUE 6.00e-81
; OTHER INFORMATION: SWISSPROT HIT: P5200, EVALUE 2.00e-04
; OTHER INFORMATION: EST_HUMAN HIT: A1138404.1, EVALUE 5.00e-76
US-10-029-386-15211

Query Match 14.9%; Score 149.8; DB 15; Length 153;
Best Local Similarity 96.7%; Pred. No. 2.5e-28;
Matches 151; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 477 CGATGTGCAATTTGTAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAATGTACC 536
DB 153 CGATGTGCAATTTGTAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAATGTACC 94

QY 537 CAGATGTATCATTCCTTGTGTCGAGGAGCGGCTCTTTTCAGGATTCAGTCAATC 596
DB 93 CAGATGTATCATTCCTTGTGTCGAGGAGCGGCTCTTTTCAGGATTCAGTCAATC 34

QY 597 TTCTCTGCTTTGTCCAGAACACATTTGACCAAGCT 629
DB 33 TTCTCTGCTTTGTCCAGAACACATTTGACCAAGCT 1

RESULT 11
US-10-221-714A-461
; Sequence 461, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of diseases associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221.714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 461
; LENGTH: 6292
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-461

Query Match 14.9%; Score 149.8; DB 13; Length 6292;
Best Local Similarity 80.6%; Pred. No. 2.8e-27;
Matches 175; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CGCAATTTAGGCTCCGGTATCTCCGGTGGAGCTCTGTTCGGGCTTAGAGACC 60

; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15211
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21.2.0
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: NT HIT: 9114756516, EVALUE 6.00e-81
; OTHER INFORMATION: SWISSPROT HIT: P5200, EVALUE 2.00e-04
; OTHER INFORMATION: EST_HUMAN HIT: A1138404.1, EVALUE 5.00e-76
US-10-029-386-15211

Query Match 14.9%; Score 149.8; DB 15; Length 153;
Best Local Similarity 96.7%; Pred. No. 2.5e-28;
Matches 151; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 477 CGATGTGCAATTTGTAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAATGTACC 536
DB 153 CGATGTGCAATTTGTAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAATGTACC 94

QY 537 CAGATGTATCATTCCTTGTGTCGAGGAGCGGCTCTTTTCAGGATTCAGTCAATC 596
DB 93 CAGATGTATCATTCCTTGTGTCGAGGAGCGGCTCTTTTCAGGATTCAGTCAATC 34

QY 597 TTCTCTGCTTTGTCCAGAACACATTTGACCAAGCT 629
DB 33 TTCTCTGCTTTGTCCAGAACACATTTGACCAAGCT 1

RESULT 12
US-10-087-192-1252
; Sequence 1252, Application US/10087192
; Publication No. US2002018258A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087.192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1252
; LENGTH: 31516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1252

Query Match 9.3%; Score 93.6; DB 13; Length 31516;
Best Local Similarity 69.7%; Pred. No. 2.1e-12;
Matches 145; Conservative 0; Mismatches 54; Indels 9; Gaps 1;

QY 184 GGGGCTGGAGCAGTAAGATGCGCGCCAGAGCGGTTTTTCTGGCAATTTCTGCCAGCTGC 243
DB 16031 GTGCAACAGAGGCTGCAGTGTGCGCCAGAGTGGCTTTCTGGCTTTCGCCGCCAGCCAC 16090

QY 244 TCCAGCCAGGCTCA-----TGAAGGAGGAGTCCCTGTGCTGAGCTGGAGGTTGG 294
DB 16091 TCTGAGAAAGGCTGGAGGAGAAAGTGAAGGAGGAGTCACTGTGCTGAGGTTGG 16150

QY 295 AGCCTGGAAGACGCGACAGCTCTGTGCTTCACTTCTGAGGTTGGCAGCCAGCTGATG 354
DB 16151 AGCTTGCAGATGGACGCTCTGTGCTCACTCTTACAGTTGTGGCGGTGACACGCCAGG 16210

QY 355 GAGACGGCAGCTCAACAGGACCAATAGG 382
DB 16211 GAGACTTCAGGCTGACTGGAGTAGAAGG 16238

RESULT 13
US-10-027-128026/c
; Sequence 128026, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```



```
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128026
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128026

Query Match      8.8%; Score 88.4; DB 13; Length 407;
Best Local Similarity 98.9%; Pred. No. 2.8e-12;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 471 TTACCCGAGTGTGCAATTTTGAAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAA 530
Db 90 TTTACCCGAGTGTGCAATTTTGAAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAA 31
```

```
QY 531 TGTACCCAGATGTATCATTTATCCCTTGCT 560
Db 30 TGTACCCAGATGTATCATTTATCCCTTGCT 1
```

```
RESULT 14
US-10-027-632-128027/c
; Sequence 128027, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128027
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128027
```

```
Query Match      8.8%; Score 88.4; DB 13; Length 407;
Best Local Similarity 98.9%; Pred. No. 2.8e-12;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 471 TTACCCGAGTGTGCAATTTTGAAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAA 530
Db 90 TTTACCCGAGTGTGCAATTTTGAAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAA 31
```

```
QY 531 TGTACCCAGATGTATCATTTATCCCTTGCT 560
Db 30 TGTACCCAGATGTATCATTTATCCCTTGCT 1
```

```
RESULT 15
US-10-027-632-128028/c
; Sequence 128028, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128028
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128028
```

```
Query Match      8.8%; Score 88.4; DB 13; Length 407;
Best Local Similarity 98.9%; Pred. No. 2.8e-12;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 471 TTACCCGAGTGTGCAATTTTGAAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAA 530
Db 90 TTTACCCGAGTGTGCAATTTTGAAGCACTTTGGAGCCCACTATCAAAATGCTGTGAAGAGAAA 31

QY 531 TGTACCCAGATGTATCATTTATCCCTTGCT 560
Db 30 TGTACCCAGATGTATCATTTATCCCTTGCT 1
```

Search completed: July 9, 2004, 12:21:50  
Job time : 697 secs

**This Page Blank (uspto)**

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 03:21:18 ; Search time 4192 Seconds  
(without alignments)  
7152.106 Million cell updates/sec

Title: US-10-081-108-1  
Perfect score: 1004  
Sequence: 1 CGCAATTAGGCTCCGG.....ATCTTTACACTAAAAGCC 1004

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0  
Searched: 27513289 segs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	EST:
1: em_estba:*	
2: em_estum:*	
3: em_estin:*	
4: em_estmu:*	
5: em_estov:*	
6: em_estpl:*	
7: em_estic:*	
8: em_estc:*	
9: gb_estl:*	
10: gb_est2:*	
11: gb_est3:*	
12: gb_est4:*	
13: gb_est5:*	
14: gb_est6:*	
15: em_estfun:*	
16: em_estom:*	
17: em_gss_hum:*	
18: em_gss_inv:*	
19: em_gss_pln:*	
20: em_gss_vrt:*	
21: em_gss_fun:*	
22: em_gss_man:*	
23: em_gss_mus:*	
24: em_gss_pro:*	
25: em_gss_rod:*	
26: em_gss_phg:*	
27: em_gss_vrl:*	
28: gb_gss1:*	
29: gb_gss2:*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	923.8	92.0	2037	11	BC017024 Homo sapi
2	450.8	44.9	732	14	CA4233957 UI-H-F81
3	349.6	34.8	699	10	BE990756 60143191
4	276.4	27.5	961	12	BG501498 602548084

5	245.8	24.5	300	9	AU099705
6	169	16.8	940	12	EM449877
7	158.4	15.8	345	9	AI138404
8	158.4	15.8	385	13	BX282448
9	155	15.4	626	29	AG092191
10	149	14.8	160	14	H55516
11	146.4	14.6	390	9	AA459832
12	141.2	14.1	533	28	CC325294
13	141.2	14.1	560	28	CC325314
14	141.2	14.1	571	28	CC325312
15	141.2	14.1	2847	11	AK054270
16	141.2	14.1	4038	11	AK087368
17	141.2	14.1	4546	11	AK044828
18	141	14.0	202	10	BE463879
19	138.2	13.8	870	12	BG218883
20	135.4	13.5	707	10	BF723805
21	118.6	11.8	281	10	BE929018
22	102.8	10.2	736	12	BJ636576
23	95.4	9.5	235	10	BF898197
24	91.8	9.1	467	14	N78369
25	87.4	8.7	743	29	AG173554
26	87.4	8.7	616	29	AG157660
27	87.2	8.7	312	14	CD000138
28	87.2	8.7	313	28	AQ567899
29	87	8.7	337	10	BF727200
30	87	8.7	426	10	AW172727
31	87	8.7	511	9	AA424842
32	87	8.7	538	9	AU146660
33	87	8.7	690	14	CA430951
34	87	8.7	722	14	CA425559
35	86.8	8.6	724	14	CA442190
36	85.6	8.5	339	12	BG983411
37	85.6	8.5	411	9	AA218727
38	85.6	8.5	483	9	AA223258
39	85.6	8.5	494	28	AQ322524
40	85.6	8.5	500	28	AQ787637
41	85.6	8.5	538	28	AQ262907
42	85.6	8.5	546	28	AQ354790
43	85.6	8.5	582	9	AA176147
44	85.6	8.5	714	12	BG427174
45	85.4	8.5	681	10	BE738289

ALIGNMENTS

RESULT 1	BC017024	2037 bp	mRNA	linear	HTC 18-JUN-2003
LOCUS	Homo sapiens cDNA clone IMAGE:3916415, containing frame-shift errors				
DEFINITION	BC017024	Homo sapiens (human)			
ACCESSION	BC017024				
VERSION	BC017024.1	GI:16877541			
KEYWORDS	HTC.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2037)	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
AUTHORS					

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)  
 23288257  
 12477932  
 2 (bases 1 to 2037)  
 Direct Submission  
 Strausberg, R.  
 Submitted (05-NOV-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC) Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
 Tissue Procurement: ATCC/DCTD/DTMP  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 22 Row: 0 Column: 15  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4557346  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers  
 1. 2037  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3916415"  
 /tissue\_type="Skin, melanotic melanoma."  
 /clone\_fld="NIH\_MGC\_72"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 ORIGIN  
 Query Match 92.0%; Score 923.8; DB 11; Length 2037;  
 Best Local Similarity 94.4%; Pred. No. 3.6e-220;  
 Matches 1002; Conservative 0; Mismatches 2; Indels 57; Gaps 2;  
 QY 1 GCCCAATTAGGCTCTCGGATATCCCGCTGAGCTCTGTGTCGGCTTACAGACC 60  
 Db 25 CGCCAAATTAGGCTCTCGGATATCCCGCTGAGCTCTGTGTCGGCTTACAGACC 84  
 QY 61 AGGAGAGGAGCTGAGAGCTGAGAGCTGACCGCTACACCGTGGCTGCTCCTGATG 120  
 Db 85 AGGAGAGGAGCTGAGAGCTGAGAGCTGACCGCTGACACCGTGGCTGCTCCTGATG 144  
 QY 121 GTGTGGCAACAGAGATGGCAGCGAGCTGAGAGTGTAGGAGGCGCGCTGAGCGTAGG 180  
 Db 145 GTGTGGCAACAGAGATGGCAGCGAGCTGAGAGTGTAGGAGGCGCGCTGAGCGTAGG 204  
 QY 191 AGTGGGCTGAGCAGTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGC 240  
 Db 205 AGTGGGCTGAGCAGTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGC 264  
 QY 241 TGCTCCAGCCAGCTGATGAGAGAGAGTCCCTGCTGAGTGGAGTGGAGCTGAGAGCTG 300  
 Db 265 TGCTCCAGCCAGCTGATGAGAGAGAGTCCCTGCTGAGTGGAGTGGAGCTGAGAGCTG 324  
 QY 301 AAGCGGCACAGCTCTGTGCTTCTCTCTGAGTGTGGAGCCAGCCAGCTGAGAGCAG 360

Db 325 AAGACGGCACAGCTCTGTGCTTCTATCTCTGAGGTTGGCAGCCAGGTGATGGAGACG 384  
 QY 361 GCAGCTCAACAGGACCAATAGGAG-----GAGATGGAGTTTCACT 400  
 Db 385 GCAGCTCAACAGGACCAATAGGAGGTTAACCGTGGAGCCCAAGTGAATGGAGTTTCACT 444  
 QY 401 GTGTCAAGCAGGATGGTCTCGATCTCTGACCTCGTACCTCGTATCCGCCCGCTTGGCTTCCAA 460  
 Db 445 GTGTCAAGCAGGATGGTCTCGATCTCTGACCTCGTATCCGCCCGCTTGGCTTCCAA 504  
 QY 461 AGTCCCGAGATTA-----CAGCGATGTG 483  
 Db 505 AGTCCCGAGATTAAGGTTTATGAAGAGATTTCTCTCTTCCCTTTTTCAGCGATGTG 564  
 QY 484 CATTTTGAAGCACTTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGACCCAGATGT 543  
 Db 565 CATTTTGAAGCACTTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGACCCAGATGT 624  
 QY 544 ATCATTTCTGTGTGTCAGGAGCGGCTCTCTTTCAGGATTTTCAGTCACATCTTCTGTC 603  
 Db 625 ATCATTTCTGTGTGTCAGGAGCGGCTCTCTTTCAGGATTTTCAGTCACATCTTCTGTC 684  
 QY 604 TTTGTCCAGACACATTGACCAAGCTCTCTGAAAGATGTAAGTTTACTACGATAGACTTT 663  
 Db 685 TTTGTCCAGACACATTGACCAAGCTCTCTGAAAGATGTAAGTTTACTACGATAGACTTT 744  
 QY 664 TAACTTCAACCAATGTATTACTGAAATTAACAAATGTTGTAATTTCCCTGAGTGTAT 723  
 Db 745 TAACTTCAACCAATGTATTACTGAAATTAACAAATGTTGTAATTTCCCTGAGTGTAT 804  
 QY 724 TCTACTTGTATAAAGTATAATACATATCAATTAATTTCTGAGGATCATTTGCCAGA 783  
 Db 805 TCTACTTGTATAAAGTATAATACATATCAATTAATTTCTGAGGATCATTTGCCAGA 864  
 QY 784 GATTTCTGGGAGGAGAAATGTTATCAACGGTTTCATTGAAATTAATTCCTCAAAAGTTATT 843  
 Db 865 GATTTCTGGGAGGAGAAATGTTATCAACGGTTTCATTGAAATTAATTCCTCAAAAGTTATT 924  
 QY 844 TCTCAGAAAAATCAAAATAAAGTTTGCATGTTTTTATTCTTAAACATTTTAAACCA 903  
 Db 925 TCTCAGAAAAATCAAAATAAAGTTTGCATGTTTTTATTCTTAAACATTTTAAACCA 984  
 QY 904 CTGTAGATGATGTAATAGGAGCTGTGACGATTTCTGACATATATCTATAAAATTTATTA 963  
 Db 985 CTGTAGATGATGTAATAGGAGCTGTGACGATTTCTGACATATATCTATAAAATTTATTA 1044  
 QY 964 AAAATCAATCAGTATCAACATCTTTTACACTAAAAGCC 1004  
 Db 1045 AAAATCAATCAGTATCAACATCTTTTACACTAAAAGCC 1085

RESULT 2  
 CA423957/c  
 LOCUS  
 DEFINITION  
 UI-H-FE1-bed-p-04-0-UI.s1 NCI CGAP FE1 Homo sapiens cDNA clone  
 UI-H-FE1-bed-p-04-0-UI 3', mRNA sequence.  
 ACCESSION  
 CA423957  
 VERSION  
 CA423957.1 GI:24786683  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 732)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
 Tissue Procurement: James Martin  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, bento-soares@iowa.edu  
 The following repetitive elements were found in this cDNA  
 sequence: 391-493, >ALU 493-590, >ALU  
 Seq primer: M13 FORWARD  
 POLYA=Yes

# FEATURES

Location/Qualifiers  
 1..732

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-PE1-bed-p-04-0-UI"  
 /tissue\_type="Cell lines"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP FE1"  
 /note="Organ: Chondrosarcoma; Vector: pT73-Pac  
 (Pharmacia) with a modified polylinker; Site 1:  
 Site 2: Not I; NCI CGAP FE1 is a normalized cDNA library  
 derived from a pool of mRNA obtained from 3 cell lines  
 from grade II chondrosarcoma tissues. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (d)18 tail. The sequence tag for this library is  
 CGCTACGGAC. The cell lines were provided by Dr James  
 Martin from the University of Iowa.  
 TAG TISSUE=Human grade 2 chondrosarcoma cell line pool  
 TAG LIB=UI-H-PE1  
 TAG\_SEQ=CGCTACGGAC"

# ORIGIN

Query Match 44.9%; Score 450.8; DB 14; Length 732;  
 Best Local Similarity 81.5%; Pred. No. 8.1e-102;  
 Matches 595; Conservative 0; Mismatches 27; Indels 108; Gaps 2;

QY 243 CTCACACCCAGGCTGATGAGGAGGATCCCTGTGTGAGCTGGAGGTTGGAGCCTGAA 302  
 Db 732 CTCACACCCAGGCTGATGAGGAGGATCCCTGTGTGAGCTGGAGGTTGGAGCCTGAA 673  
 QY 303 GACGGCACAGCTCTGTCTTCATCTCTGAGGTTGTGGACCCACGCTGATGAGACGGC 362  
 Db 672 GACGGCACAGCTCTGTCTTCATCTCTGAGGTTGTGGACCCACGCTGATGAGACGGC 613  
 QY 363 AGTCAACAGAGCAATAGAGGAGATGGAGTTTCACTGTGTGAGCCAGGATGGTCTCGA 422  
 Db 612 AGTCAACAGAGCAATAGAGGAGACGGCATCTCTCGATGTTGCCCGGCTGATCTCGA 553  
 QY 423 TCTCCTGACCTC--GTGATCGCCCGCTTGGCTTCCAA-- 461  
 Db 552 GTCTCTGGGTTGAGGTGATCGCCCGCTCAGGCTCCCAAAAGTGTGGGATATAGGCG 493  
 QY 462 ----- 461  
 Db 492 CTGGAGTGCAATGACACGATCTCGGCTCACTGACCCGATCCCTCTGCTCCAGGTTCAA 433  
 QY 462 -----GTCCAGAGATTACAGGATGTGCATTTTGTAAAG 494  
 Db 432 GCGATTCTCTGCTCAACTCCCGAGTAGCTCGGATTACAGGATGTGCATTTTGTAAAG 373  
 QY 495 CACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGATGTATCATTTCT 554  
 Db 372 CACTTTGGAGCCACTATCAATGCTGTGAAGAGAAATGTACCCAGATGTATCATTTCT 313  
 QY 555 TGTGCTGCAGAGCCGGCTCCTTTTCAGGATTTCAGTCAATCTCTCTGCTTTGTCCAGAA 614  
 Db 312 TGTGCTGCAGAGCCGGCTCCTTTTCAGGATTTCAGTCAATCTCTCTGCTTTGTCCAGAA 253

QY 615 CACATTGACCAAGCTCTTGAAGATGTAAGTTTACTAGCATAGACTTTTAACTTCAAC 674  
 Db 252 CACATTGACCAAGCTCTTGAAGATGTAAGTTTACTAGCATAGACTTTTAACTTCAAC 193  
 QY 675 CAATCTATTACTGAAGAATACAAATGTTGTAATTCCTGAGTGTATTCTTACTGTAT 734  
 Db 192 CAATGTATTACTGAAGAATACAAATGTTGTAATTCCTGAGTGTATTCTTACTGTAT 133  
 QY 735 TAAAGGTAATTAATACATAATCATTAATAATCTGAGGATCAATTCAGAGATTTGTGGG 794  
 Db 132 TAAAGGTAATTAATACATAATCATTAATAATCTGAGGATCAATTCAGAGATTTGTGGG 73  
 QY 795 AGGGAATGTTATCAACGGTTTCATTGAAATTAATTCAGAAAGTTATTCTTCAGAAA 854  
 Db 72 AGGGAATGTTATCAACGGTTTCATTGAAATTAATTCAGAAAGTTATTCTTCAGAAA 13  
 QY 855 ATCAATAAAA 864  
 Db 12 AAAAAAAAAA 3

RESULT 3  
 BE90756  
 LOCUS 601431191F1 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:3916415 5',  
 DEFINITION BE90756 699 bp mRNA linear EST 20-OCT-2000  
 BE90756 mRNA sequence.  
 BE90756.1 GI:10349397  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTP  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLN9741 row: e column: 24  
 High quality sequence stop: 407.  
 Location/Qualifiers  
 1..699  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3916415"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 72"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."

FEATURES  
 source

ORIGIN

Query Match 34.8%; Score 349.6; DB 10; Length 699;  
 Best Local Similarity 89.4%; Pred. No. 1.8e-76;  
 Matches 454; Conservative 0; Mismatches 24; Indels 30; Gaps 6;

QY 1 CGCCAAATTAGGCTCCGGTATCTCCGGTATCTCCGGTGAAGTGTCTGTTCGCCGTTAGAGACC 60  
 Db 4 CGCCAAATTAGGCTCCGGTATCTCCGGTATCTCCGGTGAAGTGTCTGTTCGCCGTTAGAGACC 63  
 QY 61 AGGGAAGGGGAGCTGAGGCTGTAAACACCGTGGCTCGTCTCACTCTGATG 120

64	AGGAGAGGGGGAGCTGGAGGCTGGAGCGCTGTAAACACCGTGCGTCTCACTCTGGATG	123
121	GTGCTGGCAAC-AGAGATGCACGGCAGCTGGAGTGTTAGGAGGCGCGCTTGACGGTAG	179
124	GTGTGGGCACAAGAGATGCGCAGCTGGAGTGTAGGAGGCGCGCTTGACGGTAG	183
180	GAGTGGGCTGGACAGTA-AGATGGCGGCAGAGCGGTTTTCTGGCAATTGTCTGCCCA	238
184	GAGTGGGCTGGACAGTACAGATGGCGCCAGAGCGGTTTTCTGGCATGCTCTGCCCA	243
239	GCTGCTCAAGCCAGGCTGATGAAGAGAGAGTCCCTGTGTGAGCTCGAGGTTGAGACC	298
244	GCTGCTCAAGCCAGGCTGATGAAGAGAGAGTCCCTGTGTGAGCTCGAGGTTGAGACC	303
299	TGAAGACGGCACAGCTCTGTGCTTCATCTCTGAGGTTGTGCACGCCACGGTGATGAGA	358
304	TGAAGACGGCACAGCTCTGCGCTTCATCTCTGAGGTTGTGCACGCCACGGTGATGAGA	363
359	CGGCAGC-----TCACAGGAGCAATAGCAGGAGATGGAGT	394
364	CGGCAGCCTCAACACGGACCACTTACGAGGTTAACGTTGGAGGCCACTGAGATGCCAGT	423
395	TTCACTGTGTGAG-CCAGGATGTTCTCG-ATCTCTGACCTCGTGAT--CGGCCCGCGCTT	450
424	CTCACTGTGTGAGCCCAAGGATGTTCTGAATCTCTGACCTCGTGATTTCTGCCCGCGCTT	483
451	GGCCTTCCAAAGTGCCGAGATTACGGC	478
484	GGCCTTCCAAAGTGGCGAGATTACGCG	511

[illegible]

ACCESSION BG501498  
 BG501498.1 GI:13463015  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 961)  
 N1H-MGC <http://mgc.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg Ph.D.

3' adaptor sequence:  
5'-ATTCTAGCCGAGCGCCGACATG-dT(30)BN-3' (where B = A, C, G or N = A, C, G, or T). Average insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH MGC Library."

ORIGIN	Query Match	27.5%; Score 276.4; DB 12; Length 961;
	Best Local Similarity	92.3%; Pred. No. 4.1e-58;
	Matches 310; Conservative	0; Mismatches 6; Indels 20; Gaps 1;
QY	1	CGCCAAATTAGGGTCTCCGGTATCTCCCGCTGAGCTCTGTTCCCGGCTTTAGAGGACC 60
DB	44	CGCCAGATTAGGGTCTCTGGTATCTCCCGCTGAGCTGCTCTGTTTCCGGCTTTAGAGGACC 103
QY	61	AGGAGAGAGGGGAGCTGGAGGCTTGGAGCCTGTGAACAACCGTGGCTCGTCTCACTCTGGATG 120
DB	104	AGGAGAGAGGGGAGTTGGAGGCTTGGAGGCTGTGAACAACCGTGGCTCGTCTCGCTCTGGATG 163
QY	121	GTGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTTAGAGGGGGGCGCTGAGCGGTAGG 180
DB	164	GTGTGGCAACAGAGATGGCAGCGCAGCTGGAGTGTTAGAGGGGGCGGCTGAGCGGTAGG 223
QY	181	AGTGGGCTGGAGCAGTAAAGATGGCGGCCAGAC-----GGTTTT 220
DB	224	AGTGGGCTGGAGCAGTAAAGATGGCGGCCGAGCAGTAAAGTGGCGGCTGGAGTGGTTTT 283
QY	221	TCTGGCATTTGTCGCCCAGCTGTCTCAAGCCAGGCTGATGAAGGAGGAGTCCCCCTGTGGT 280
DB	284	TCTGGCATTTGTCGCCCAGCTGTCTCAAGCCAGGCTGATGAAGGAGGAGTCCCCCTGTGGT 343
QY	281	GAGCTGGAGTTGGAGCCTGAAGCAGGCAACAGCTCT 316
DB	344	GAGCTGGAGTTGGAGCCTGAAGTGCACAGCTCT 379

RESULT 5						EST 05-APR-2001
AU099705	Sugano Homo sapiens cDNA library	linear				
LOCUS	KAT00969 similar to Human B melanoma antigen (BAGE) mRNA,	mRNA	300 bp			
DEFINITION	sequence.					
ACCESSION	AU099705					
VERSION	AU099705.1	GI:13550834				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa;	Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria;	Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 300)					
AUTHORS	Suzuki,Y., Tsunoda,T., Taiba,H., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo,K., Suyama,A. and Sugano.S.					
TITLE	In silico mapping of the 5'-ends of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries constructed by					

```

plate: LLCML483 row: n column: 22
High quality sequence stop: 689.
Location/Qualifiers
1. 961
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4670277"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_60"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech) ;
Site: 1: Sfil (ggcgctctggcc) ; Site: 2: Sfil
(ggcattatggcc) ; Double-stranded cDNA was prepared f
cell line RNA. 5' and 3' adaptors were used in clon
follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3

```

```
/clone="KAT00969"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match
Best Local Similarity 24.5%; Score 245.8; DB 9; Length 300;
Matches 267; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 17 CCGTATCTCCGCTGAGCTGCTCTGTTCCCGGCTTAGAGGACCAAGGGGAGCT 76
Db 17 CCTGATCTCCCGCTGAGCTGCTCTAGTCCCGGCTTAGAGGACCAAGGGGAGCT 60
QY 77 GGAGCTGGAGCTGTAACACCGTGGCTCTGCTCACTCTGATGGTGGTGGCAACAGAGA 136
Db 77 GGAGCTGGAGCTGTAACACCGTGGCTCTGCTCACTCTGATGGTGGTGGCAACAGAGA 120
QY 137 TGGCAGCGCAGCTGAGTCTTAGGAGGGGGCGCTTACCTCTGATGGTGGTGGCAACAGAGA 196
Db 137 TGGCAGCGCAGCTGAGTCTTAGGAGGGGGCGCTTACCTCTGATGGTGGTGGCAACAGAGA 180
QY 197 TAAGATGGCGGCGCAGAGCGGT 217
Db 181 TAGGAGGGGGCTGGAGCAGT 201

RESULT 7
AII38404 345 bp mRNA linear EST 28-OCT-1998
LOCUS qd84b05.x1 Soares testis NHT Homo sapiens cDNA clone IMAGE:1736145
DEFINITION 3', similar to TR:O14686 O14686 ALR.; mRNA sequence.
ACCESSION AII38404
VERSION AII38404.1 GI:3644376
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 345)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 428 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
Location/Qualifiers
1. 345
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1736145"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo (dT)
primer [5].
TGTTACCAATCTGAAGTGGGCGGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cots, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 15.8%; Score 158.4; DB 9; Length 345;
Best Local Similarity 99.4%; Pred. No. 1.1e-28;
Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
/clone="KAT00969"
/clone_lib="Sugano Homo sapiens cDNA library"

ORIGIN
Query Match
Best Local Similarity 24.5%; Score 245.8; DB 9; Length 300;
Matches 267; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 17 CCGTATCTCCGCTGAGCTGCTCTGTTCCCGGCTTAGAGGACCAAGGGGAGCT 76
Db 17 CCTGATCTCCCGCTGAGCTGCTCTAGTCCCGGCTTAGAGGACCAAGGGGAGCT 60
QY 77 GGAGCTGGAGCTGTAACACCGTGGCTCTGCTCACTCTGATGGTGGTGGCAACAGAGA 136
Db 77 GGAGCTGGAGCTGTAACACCGTGGCTCTGCTCACTCTGATGGTGGTGGCAACAGAGA 120
QY 137 TGGCAGCGCAGCTGAGTCTTAGGAGGGGGCGCTTACCTCTGATGGTGGTGGCAACAGAGA 196
Db 137 TGGCAGCGCAGCTGAGTCTTAGGAGGGGGCGCTTACCTCTGATGGTGGTGGCAACAGAGA 180
QY 197 TAAGATGGCGGCGCAGAGCGGT 217
Db 181 TAGGAGGGGGCTGGAGCAGT 201

RESULT 6
BM449877 940 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6393056 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528612
DEFINITION 5', mRNA sequence.
ACCESSION BM449877
VERSION BM449877.1 GI:18498917
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12205 row: 1 column: 21
High quality sequence stop: 423.
Location/Qualifiers
1. 940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5528612"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 16.8%; Score 169; DB 12; Length 940;
Best Local Similarity 90.0%; Pred. No. 3.2e-31;
```

471 TTACAGCGATGTCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530  
 149 TTTACAGCGATGTCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 208  
 531 TGTACCCAGATGATCATTTATCTTGTGCTGCAGGAGCGGCTCTTTTCAGGATTTTCAGT 590  
 209 TGTACCCAGATGATCATTTATCTTGTGCTGCAGGAGCGGCTCTTTTCAGGATTTTCAGT 268  
 591 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTC 630  
 269 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTC 308

RESULT 8  
 BX282448/c  
 LOCUS BX282448 Soares testis NHT Homo sapiens cDNA clone IMAGE9981104411  
 DEFINITION ; IMAGE:1736145, mRNA sequence.  
 ACCESSION BX282448  
 VERSION BX282448.1 GI:28612995  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 385)  
 Ebert,L., Hell,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,  
 Radlof,U., Schneider,D. and Korn,B.  
 Human Unigeneset - RZPD3  
 Unpublished (2003)  
 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD; IMAGE9981104411.  
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
 bin/showlib.pl/cgi/response?libNo=972 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTACACAGGAACAGCTATGAC.

FEATURES  
 source  
 1..385  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE9981104411 ; IMAGE:1736145"  
 /sex="male"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares testis NHT"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5].  
 TGTACCAATGTAAGTGGAGCGCGCCCAATTTTCTTTT 3'.  
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
 Query Match 15.8%; Score 158.4; DB 13; Length 385;  
 Best Local Similarity 99.4%; Pred. No. 1.2e-28;  
 Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 471 TTACAGCGATGTCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 530

229 TTTACAGCGATGTCATTTTGTAGCACTTTGGAGCCACTATCAATGCTGTGAAGAGAA 170  
 531 TGTACCCAGATGATCATTTATCTTGTGCTGCAGGAGCGGCTCTTTTCAGGATTTTCAGT 590  
 169 TGTACCCAGATGATCATTTATCTTGTGCTGCAGGAGCGGCTCTTTTCAGGATTTTCAGT 110  
 591 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTC 630  
 109 CACATCTTCTGCTTTGTCCAGAACACATTTGACCAAGCTC 70

RESULT 9  
 AG092191/c  
 LOCUS Pan troglodytes DNA, clone: PTB-092E12.F, genomic survey sequence.  
 DEFINITION AG092191  
 ACCESSION AG092191.1 GI:16643993  
 VERSION  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 1  
 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
 Totoki,Y., Watanabe,H. and Sakaki,Y.  
 BAC end sequences of Library PTB  
 Unpublished  
 2 (bases 1 to 626)  
 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
 Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Direct Submision  
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail:chimbases@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,  
 Tel:81-45-503-9111, Fax:81-45-503-9370)  
 Clones are derived from the chimpanzee BAC library PTB This BAC end  
 was generated during the R&D process and may have higher chance of  
 clone tracking errors.  
 PRIMERS  
 Sequencing: -21M13  
 LIBRARY  
 Vector : pKS145  
 R.Site 1 : SacI  
 R.Site 2 : SacI.  
 Location/Qualifiers  
 1..626  
 /organism="pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-092E12.F"  
 /sex="male"  
 /cell\_type="lymphoblast"  
 /clone\_lib="PTB Chimpanzee Male BAC Library"

FEATURES  
 source  
 1..626  
 /organism="pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-092E12.F"  
 /sex="male"  
 /cell\_type="lymphoblast"  
 /clone\_lib="PTB Chimpanzee Male BAC Library"

ORIGIN  
 Query Match 15.4%; Score 155; DB 29; Length 626;  
 Best Local Similarity 94.2%; Pred. No. 9.4e-28;  
 Matches 161; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 215 GGTCTTTCTGGCATTTGCTGCCAGCTCTCAAGCCAGGCTGATGAAGAGGAGTCCCC 274  
 397 GGTCTTTCTGGCATTTGCTGCCAGCTCTCAAGCCAGGCTGATGAAGAGGAGTCCCC 338  
 275 TGTGTGAGCTGGAGCTTGGAGCTGAAGAGCGGACAGCTCTGTCTTCATCTTCTGAGG 334  
 337 TGTGTGAGCTGGAGCTTGTAGCTTGAAGATGGCACAGCTCTGTGATTCATCTCTCGG 278  
 335 TTGTGGCCAGCCAGCTGATGGAGCGGAGCTCAACAGGAGCAATAGGAGG 385  
 277 TTGTGGCGCCACGGTGTGAGAGCTGCAGCTCAACAGGAGTGTAGGAGG 227





COMMENT Contact: BayGenomics  
Bay Area Functional Genomics Consortium (BayGenomics)  
Email: info@baygenomics.ucsf.edu  
Sequence tag generated by 5' RACE of total RNA from gene trap ES  
cell line. ES cell lines harboring insertion mutation of target  
gene are available upon request from BayGenomics. Annotation  
information available from  
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=  
CELL\_LINE&KEY=XM083  
Class: Gene Trap.  
Location/Qualifiers  
1. 533  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129 OLA"  
/db\_xref="taxon:10090"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
/note="Vector: pGT0Lxf"

FEATURES source  
Query Match 14.1%; Score 141.2; DB 28; Length 533;  
Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
Matches 155; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 467 GAGATTACAGCGATGTCATTTTGAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGA 526  
|||  
DB 103 GAGCACAGACGATGTGCATTTTGCACCAACACCTTGGAGCCACTATCAAAATGCTGTGAAGA 162  
|||  
QY 527 GAAATGTACCCAGATGATCATTTCTGTGTCAGAGCGCGCTCCTTTTCAGGATTT 586  
|||  
DB 163 GAAGTGATCCAGATGATCAATTCATCCATGCTGCGAGGGCTGGACCTTTTCAGGACTT 222  
|||  
QY 587 CAGTCACATCTTCTGCTTGTCCAGAACACATTCAGCAAGCTCCTGAAAGATGTAAG 644  
|||  
DB 223 TAGTCATCTTCTCTCTCTGTCAGAACACATCGACCAAGCTCTGAAAGATCAAG 280  
|||

RESULT 13  
CC325314 560 bp mRNA linear GSS 14-MAY-2003  
LOCUS XM083 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA  
DEFINITION sequence.  
ACCESSION CC325314 GI:30719372  
VERSION CC325314  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 560)  
BayGenomics.  
http://baygenomics.ucsf.edu/  
TITLE Unpublished (2001)  
JOURNAL  
COMMENT Contact: BayGenomics  
Bay Area Functional Genomics Consortium (BayGenomics)  
Email: info@baygenomics.ucsf.edu  
Sequence tag generated by 5' RACE of total RNA from gene trap ES  
cell line. ES cell lines harboring insertion mutation of target  
gene are available upon request from BayGenomics. Annotation  
information available from  
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=  
CELL\_LINE&KEY=XM083  
Class: Gene Trap.  
Location/Qualifiers  
1. 560  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129 OLA"  
/db\_xref="taxon:10090"  
/sex="Male"  
/cell\_type="Embryonic stem cell"

FEATURES source  
Query Match 14.1%; Score 141.2; DB 28; Length 571;  
Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
Matches 155; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 467 GAGATTACAGCGATGTCATTTTGAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGA 526  
|||  
DB 112 GAGCACAGACGATGTGCATTTTGCACCAACACCTTGGAGCCACTATCAAAATGCTGTGAAGA 171  
|||  
QY 527 GAAATGTACCCAGATGATCATTTCTGTGTCAGAGCGCGCTCCTTTTCAGGATTT 586  
|||  
DB 172 GAAGTGATCCAGATGATCAATTCATCCATGCTGCGAGGGCTGGACCTTTTCAGGACTT 231  
|||  
QY 587 CAGTCACATCTTCTGCTTGTCCAGAACACATTCAGCAAGCTCCTGAAAGATGTAAG 644  
|||

ORIGIN  
Query Match 14.1%; Score 141.2; DB 28; Length 560;  
Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
Matches 155; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 467 GAGATTACAGCGATGTCATTTTGAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGA 526  
|||  
DB 101 GAGCACAGACGATGTGCATTTTGCACCAACACCTTGGAGCCACTATCAAAATGCTGTGAAGA 160  
|||  
QY 527 GAAATGTACCCAGATGATCATTTCTGTGTCAGAGCGCGCTCCTTTTCAGGATTT 586  
|||  
DB 161 GAAGTGATCCAGATGATCAATTCATCCATGCTGCGAGGGCTGGACCTTTTCAGGACTT 220  
|||  
QY 587 CAGTCACATCTTCTGCTTGTCCAGAACACATTCAGCAAGCTCCTGAAAGATGTAAG 644  
|||  
DB 221 TAGTCATCTTCTCTCTCTGTCAGAACACATCGACCAAGCTCTGAAAGATCAAG 278  
|||

RESULT 14  
CC325312 571 bp mRNA linear GSS 14-MAY-2003  
LOCUS XM085 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA  
DEFINITION sequence.  
ACCESSION CC325312 GI:30719370  
VERSION CC325312  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 571)  
BayGenomics.  
http://baygenomics.ucsf.edu/  
TITLE Unpublished (2001)  
JOURNAL  
COMMENT Contact: BayGenomics  
Bay Area Functional Genomics Consortium (BayGenomics)  
Email: info@baygenomics.ucsf.edu  
Sequence tag generated by 5' RACE of total RNA from gene trap ES  
cell line. ES cell lines harboring insertion mutation of target  
gene are available upon request from BayGenomics. Annotation  
information available from  
http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=  
CELL\_LINE&KEY=XM085  
Class: Gene Trap.  
Location/Qualifiers  
1. 571  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129 OLA"  
/db\_xref="taxon:10090"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
/note="Vector: pGT0Lxf"

ORIGIN  
Query Match 14.1%; Score 141.2; DB 28; Length 571;  
Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
Matches 155; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 467 GAGATTACAGCGATGTCATTTTGAAGCACTTTGGAGCCACTATCAAAATGCTGTGAAGA 526  
|||  
DB 112 GAGCACAGACGATGTGCATTTTGCACCAACACCTTGGAGCCACTATCAAAATGCTGTGAAGA 171  
|||  
QY 527 GAAATGTACCCAGATGATCATTTCTGTGTCAGAGCGCGCTCCTTTTCAGGATTT 586  
|||  
DB 172 GAAGTGATCCAGATGATCAATTCATCCATGCTGCGAGGGCTGGACCTTTTCAGGACTT 231  
|||  
QY 587 CAGTCACATCTTCTGCTTGTCCAGAACACATTCAGCAAGCTCCTGAAAGATGTAAG 644  
|||

Db 232 TAGTCACCTTCTCTCTCTCTGTCCAGAACACATCGACCAAGCTCTCGAAGATCAAG 289

RESULT 15  
AK054270  
LOCUS  
DEFINITION  
AK054270 2647 bp mRNA linear HTC 20-SEP-2003  
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN  
full-length enriched library, clone:E330008K23 product:hypothetical  
HMG-I and HMG-Y DNA-binding domain (A+T-hook)/PHD-finger/DHHC-type  
Zn-finger/RING finger containing protein, full insert sequence.

AK054270  
AK054270.1 GI:26344110  
HTC; CAP trapper.  
Mus musculus (house mouse)  
SOURCE  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
10349636  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

20499374  
11042159  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, Y., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913  
11076861  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 2647)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Haragaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,

COMMENT  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .2647  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:E330008K23"  
/db\_xref="MGI:2426314"  
/db\_xref="taxon:10090"  
/clone="E330008K23"  
/sex="female"  
/tissue\_type="ovary"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days pregnant adult"  
207. .2645  
/note="unnamed protein product; hypothetical HMG-I and  
HMG-Y DNA-binding domain (A+T-hook)/PHD-finger/DHHC-type  
Zn-finger/RING finger containing protein  
InterPro|IPR000637, InterPro|IPR001965,  
InterPro|IPR001594, InterPro|IPR001841, evidence:  
putative"  
/codon\_start=1  
/protein\_id="BAC35712.1"  
/db\_xref="GI:26344111"  
/translation="MSSSEDSASQQPPPEPGAPSPAAAKRPRGRPKRG  
APFQARAKPRSRGKSTVEDSDGLTETENIVETEKEQSVDEDAETVDSSK  
QPSALQKSVSEANSLSVSGVKEISQELCAFQCYCEKSLGQDLKQPRVTPGLT  
LPWKDPSNKIDIDNSGCTKIQNYAPRKQGRKRPQCSAVSCVSVSTQACED  
QAGLQKDSLSVGLPDAIDVALFDSTGTWHRHRCVEMWSIGICQMEPLAVNDKAV  
VSGSTERCAFCKHGLACTKCECEKCTOMVHYPCAAGATFQDFSHFLLCFEHDQAP  
ERSKEDANCAVDSFGDLLDQFFCTCGQHYGMCLDIATVPLKRAQCCEKVCQN  
CKQSGEDSKMLVCDTCDGYYTFCQLQPMKSVPTNGMKCKNCRICIEGTSRQWHH  
NCLIDTCYQQDNLCPFGCGHPELQKMLHCNMCKRWYHLECDFTDQLDSQK  
EDYICMYCKHGLAEIDLPKGNVEVMEPLPTDYASGMIECTEDVEVLEQTVNKKVS  
DHQCRPGTVPDVQVVTTEBPQKSNPLESPDTVTLITSSNDKMPDLANEFAHSDPTE  
KTEMLSKGRHVCEEDQEDNEMVTENIVLPHQTIPOEDLLSDESEVACKELSPK  
SAPETAAPKALLSPHSRESLSCKEPLLTERTVQEMEQKENSEFSTGCVDFMTLAVDS  
CDKSSSCQGDXYELPAREESTFSSADLNKADVSSSTLCLDPSCDMLHGYPAFN  
SAAGSIMPTTIVT"

FEATURES  
source  
1. .2647  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:E330008K23"  
/db\_xref="MGI:2426314"  
/db\_xref="taxon:10090"  
/clone="E330008K23"  
/sex="female"  
/tissue\_type="ovary"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days pregnant adult"  
207. .2645  
/note="unnamed protein product; hypothetical HMG-I and  
HMG-Y DNA-binding domain (A+T-hook)/PHD-finger/DHHC-type  
Zn-finger/RING finger containing protein  
InterPro|IPR000637, InterPro|IPR001965,  
InterPro|IPR001594, InterPro|IPR001841, evidence:  
putative"  
/codon\_start=1  
/protein\_id="BAC35712.1"  
/db\_xref="GI:26344111"  
/translation="MSSSEDSASQQPPPEPGAPSPAAAKRPRGRPKRG  
APFQARAKPRSRGKSTVEDSDGLTETENIVETEKEQSVDEDAETVDSSK  
QPSALQKSVSEANSLSVSGVKEISQELCAFQCYCEKSLGQDLKQPRVTPGLT  
LPWKDPSNKIDIDNSGCTKIQNYAPRKQGRKRPQCSAVSCVSVSTQACED  
QAGLQKDSLSVGLPDAIDVALFDSTGTWHRHRCVEMWSIGICQMEPLAVNDKAV  
VSGSTERCAFCKHGLACTKCECEKCTOMVHYPCAAGATFQDFSHFLLCFEHDQAP  
ERSKEDANCAVDSFGDLLDQFFCTCGQHYGMCLDIATVPLKRAQCCEKVCQN  
CKQSGEDSKMLVCDTCDGYYTFCQLQPMKSVPTNGMKCKNCRICIEGTSRQWHH  
NCLIDTCYQQDNLCPFGCGHPELQKMLHCNMCKRWYHLECDFTDQLDSQK  
EDYICMYCKHGLAEIDLPKGNVEVMEPLPTDYASGMIECTEDVEVLEQTVNKKVS  
DHQCRPGTVPDVQVVTTEBPQKSNPLESPDTVTLITSSNDKMPDLANEFAHSDPTE  
KTEMLSKGRHVCEEDQEDNEMVTENIVLPHQTIPOEDLLSDESEVACKELSPK  
SAPETAAPKALLSPHSRESLSCKEPLLTERTVQEMEQKENSEFSTGCVDFMTLAVDS  
CDKSSSCQGDXYELPAREESTFSSADLNKADVSSSTLCLDPSCDMLHGYPAFN  
SAAGSIMPTTIVT"

Query Match 14.1%; Score 141.2; DB 11; Length 2647;  
Best Local Similarity 87.1%; Pred. No. 3.9e-24;  
Matches 155; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 467 GAGATTACAGGATGTGCATTTTGTAAAGCACTTTGGAGCCACTATCAAACTCTGTGAAGA 526  
DB 1043 GAGCACAGACGATGTGCATTTTCAAAACACCTTTGGAGCCACTATCAAACTCTGTGAAGA 1102  
QY 527 GAAATGTACCCAGATGATCATTTCTGTGCTGCAGGAGCCGGCTCTTCAGGATTT 586  
DB 1103 GAAGTGTACCCAGATGATCATTTCTGTGCTGCAGGAGCCGGCTCTTCAGGATTT 1162  
QY 587 CAGTCACATCTTCTCTGCTTTTCCAGAACATTTGACCAAGCTCTCTGAAGATGTAAG 644  
DB 1163 TAGTCACATCTTCTCTGCTTTTCCAGAACATTTGACCAAGCTCTCTGAAGATGTAAG 1220

Search completed: July 9, 2004, 10:07:38  
Job time : 4200 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 16:57:41 ; Search time 55 Seconds  
(without alignments)  
220.901 Million cell updates/sec

Title: US-10-081-108-2  
Perfect score: 213  
Sequence: 1 MARAVFLAQAQLQARLM.....SPVSVRLPEDEGTCALCFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	2	Aaw02152 BAGE tumo
2	213	100.0	43	5	Aau84811 Human BAG
3	213	100.0	43	5	Abb78347 Amino aci
4	213	100.0	43	7	Add25520 Binding d
5	156	73.2	30	5	Aau85010 Human BAG
6	156	73.2	3541	5	Aau85130 Human mel
7	126	59.2	30	5	Aau85009 Human BAG
8	98	46.0	22	2	Aar67808 BAGE tumo
9	98	46.0	22	2	Aay10634 Peptide a
10	98	46.0	22	5	Abg80319 MHC class
11	87	40.8	17	5	Aau85011 Human BAG
12	70	32.9	16	2	Aar67809 BAGE tumo
13	70	32.9	16	2	Aay10635 Peptide a
14	70	32.9	16	5	Abg80316 MHC class
15	63	29.6	384	4	Aab48007 Human sph
16	63	29.6	384	6	Abp71054 Human sph
17	62	29.1	293	2	Aaw88613 Secreted
18	62	29.1	293	4	Abb50380 Human sec
19	62	29.1	293	6	Abc44637 Novel hum
20	62	29.1	293	7	Abc26117 Human pro
21	62	29.1	305	4	Aam41966 Human pol
22	62	29.1	333	5	Abb90209 Human pol
23	62	29.1	368	6	Abrr2391 Human sph
24	62	29.1	394	3	Aab18659 A human r
25	62	29.1	394	3	Aay96057 Human sph

26	62	29.1	384	4	AAW40180	Aam40180 Human pol
27	62	29.1	384	4	AAE94589	Aab94589 Human pro
28	62	29.1	384	4	AAE93955	Aab93955 Human pro
29	62	29.1	384	4	AAE00324	Aae00324 Human sph
30	62	29.1	384	4	AAE07882	Aae07882 Human sph
31	62	29.1	384	5	ABG31586	Abg31586 Human pro
32	62	29.1	384	7	ADE38365	Ade38365 Human pro
33	60	28.2	552	6	ABU47220	Abu47220 Protein e
34	58	27.2	173	4	ABB71060	Abb71060 Drosophil
35	58	27.2	384	5	ABB08089	Abb08089 Human sph
36	57.5	27.0	340	6	ABU44326	Abu44326 Protein e
37	57	26.8	398	2	AAR78673	Aar78673 CD4 domai
38	57	26.8	398	2	AAR89450	Aar89450 CD4 DI-D4
39	57	26.8	399	1	AAP93010	Aap93010 Genetic c
40	57	26.8	416	3	AAB19509	Aab19509 CD4-IGM f
41	57	26.8	432	2	AAW74222	Aay74222 Epitope o
42	57	26.8	432	2	AAW30514	Aay30514 Predicted
43	57	26.8	432	6	ADA25188	Ada25188 CD4 epitc
44	57	26.8	433	2	AAW41376	Aaw41376 Human CD4
45	57	26.8	433	3	AAW54500	Aay54500 Amino aci

## ALIGNMENTS

## RESULT 1

AAW02152  
ID AAW02152 standard; protein; 43 AA.

XX AAW02152;

AC AAW02152;

DT 04-DEC-1996 (first entry)

XX BAGE tumour rejection antigen precursor.

DE BAGE tumour rejection antigen precursor.

XX BAGE; tumour rejection antigen precursor; TRAP; MHC;

KW major histocompatibility complex; HLA-Cw\*1601; melanoma; metastasis;

KW diagnosis; therapy; vaccine.

XX Homo sapiens.

XX Key

FT Peptide

FT Location/Qualifiers

FT 2..10

FT /label= Tumour\_rejection\_antigen

XX WO9625511-A1.

XX 22-AUG-1996.

XX 07-FEB-1996; 96WO-US001608.

XX 16-FEB-1995; 95US-00389360.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Boel P, Wildmann C, Boonfalleur T, Van Der Bruggen P, Coullie P;

XX Renauld J;

XX WPI; 1996-393411/39.

XX N-FSDB; AAR36382.

XX Tumour rejection antigen precursor (TRAP) and gene - useful to develop

XX prods. for diagnosis and treatment of disorders characterised by TRAP,

XX partic. melanomas.

XX Example 4; Page 11; 44pp; English.

XX The BAGE tumour rejection antigen precursor (TRAP) (AAW02152) is

XX processed to a tumour rejection antigen (TRA) (AAW02153) presented by the

XX MHC molecule HLA-Cw\*1601. The amino acid sequence of the BAGE TRAP was

XX deduced from a cDNA clone (AAT36382) derived from melanoma cell line M22-

XX MEL-43. BAGE expression was not observed in healthy adult or foetal

XX tissue, but was detected in 22% of melanoma lines examined, being partic.

```

Query Match      100.0%; Score 213; DB 5; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.8e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MAARAVFALSAQLLQARLMKESPVVSWELEPEDTALCFIF 43
db 1 MAARAVFALSAQLLQARLMKESPVVSWELEPEDTALCFIF 43

```

RESULT 4  
ADD25520  
ID ADD25520 standard; protein; 43 AA.  
XX  
XX  
AC ADD25520;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Binding domain-immunoglobulin fusion protein-associated protein #37.  
XX  
XX Binding domain; immunoglobulin; fusion protein; cytostatic;  
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
KW neuroprotective; hinge region; immunoglobulin heavy chain;  
KW CH2 constant region; CH3 constant region; IgG1;  
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX  
XX Unidentified.  
XX  
XX US2003118592-A1.  
XX  
XX 26-JUN-2003.  
XX  
XX 25-JUL-2002; 2002US-00207655.  
XX  
XX 17-JAN-2001; 2001US-0367358P.  
PR 17-JAN-2002; 2002US-00053530.  
PR 03-JUN-2002; 2002US-0385691P.  
XX  
XX (GENE-) GENE-CRAFT INC.  
XX  
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
XX WPI; 2003-801317/75.  
XX  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
PT subject having or suspected of having a malignant condition or a B-cell  
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX  
XX Disclosure; SEQ ID NO 81; 157pp; English.  
XX  
XX The invention relates to a binding domain-immunoglobulin fusion protein  
CC comprising a binding domain polypeptide that is fused to an  
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
CC polypeptide that is fused to the CH2 constant region polypeptide. The  
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
CC region polypeptide; a mutated human IgG1 immunoglobulin hinge  
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
CC immunoglobulin hinge region polypeptide contains no more than one  
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
CC polypeptide, derived from (a) having 3 or more cysteine residues; where  
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
CC capable of at least one immunological activity comprising antibody  
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
CC binding domain polypeptide is capable of specifically binding to an  
CC antigen. Also included are an isolated polynucleotide encoding the  
CC binding domain-immunoglobulin fusion protein, a recombinant expression  
CC construct comprising the polynucleotide operably linked to a promoter,  
CC a host cell transformed or transfected with a recombinant expression  
CC construct, producing the binding domain-immunoglobulin fusion protein, a  
CC pharmaceutical composition comprising the binding domain-immunoglobulin  
CC fusion protein or polynucleotide and a carrier, and treating a subject  
CC having or suspected of having a malignant condition or a B-cell disorder.

CC The binding domain-immunoglobulin fusion protein is useful for treating a  
CC subject having or suspected of having a malignant condition or a B-cell  
CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
CC sclerosis or autoimmune disease. The present sequence is a binding domain  
CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
CC sequence data for this patent formed part of the printed specification  
CC and is also available in electronic format directly from USPTO at  
CC segdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not  
CC identified the sequences in the printed specification by their SEQ ID  
CC number therefore none of the sequences can be explicitly identified.  
XX  
XX Sequence 43 AA;  
SQ  
Query Match 100.0%; Score 213; DB 7; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.8e-23;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAARAVFLALSAQLLQARLMKEESPVSVWRLEPDGTALCFIF 43  
DB 1 MAARAVFLALSAQLLQARLMKEESPVSVWRLEPDGTALCFIF 43  
RESULT 5  
AAU85010  
ID AAU85010 standard; peptide; 30 AA.  
XX  
XX AAU85010;  
XX  
XX 08-MAY-2002 (first entry)  
XX  
XX Human BAGE segment 2.  
XX  
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
KW viral infection; human immunodeficiency virus; melanoma;  
KW bacterial infection; Salmonella; Legionella; parasitic infection;  
KW Trypanosoma; Toxoplasma; Giardia.  
XX  
XX Homo sapiens.  
XX  
XX WO200190197-A1.  
XX  
XX 29-NOV-2001.  
XX  
XX 25-MAY-2001; 2001WO-AU000622.  
XX  
XX 26-MAY-2000; 2000AU-00007761.  
XX  
XX (AUSU ) UNIV AUSTRALIAN NAT.  
XX  
XX Thomson SA, Ramshaw IA;  
XX WPI; 2002-147575/19.  
XX N-PSDB; ABK36830.  
XX  
XX New synthetic polypeptides having several different segments of at least  
XX one parent polypeptide linked together differently compared to the  
XX linkage in the parent polypeptide, for inducing immune response against a  
XX pathogen or cancer.  
XX  
XX Example 3; Fig 27; 364pp; English.  
XX  
XX The invention relates to a new synthetic polypeptide (I) comprising  
XX several different segments of at least one parent polypeptide linked  
XX together in a different relationship relative to their linkage in the  
XX parent polypeptide to impede, abrogate or otherwise alter at least one  
XX function associated with the parent polypeptide and for inducing an  
XX immune response against a pathogen or cancer. Also included are a  
XX synthetic polynucleotide encoding and a computer system for designing the  
XX synthetic polypeptides. The synthetic polypeptides and polynucleotides  
XX are referred to as a Savine. The synthetic polypeptide is useful for a  
XX modulating immune responses preferably directed against a pathogen or a  
XX cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head

CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,  
 CC oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is  
 CC a peptide derived from a parent protein used to construct a savine of the  
 CC invention  
 XX  
 XX Sequence 30 AA;

Query Match 73.2%; Score 156; DB 5; Length 30;  
 Best Local Similarity 100.0%; Pred. NO. 2.5e-15;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 Db 1 LLQARLMKEESPVVSWRLEPEDGTALCFIF 30

## RESULT 6

AAU85130  
 ID AAU85130 standard; protein; 3541 AA.

XX  
 AC AAU85130;

XX  
 DT 08-MAY-2002 (first entry)

XX  
 DE Human melanoma specific savine.

XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmorella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.

XX  
 OS Homo sapiens.  
 OS Synthetic.

XX  
 FN WO200190197-A1.

XX  
 PD 29-NOV-2001.

XX  
 PF 25-MAY-2001; 2001WO-AU000622.

XX  
 PR 26-MAY-2000; 2000AU-00007761.

XX  
 PA (AUSU ) UNIV AUSTRALIAN NAT.

XX  
 PI Thomson SA, Ramshaw IA;

XX  
 DR WPI; 2002-147575/19.

XX  
 DR N-PSDB; ABK36950.

XX  
 PT New synthetic polypeptides having several different segments of at least  
 PT one parent polypeptide linked together differently compared to the  
 PT linkage in the parent polypeptide, for inducing immune response against a  
 PT pathogen or cancer.

XX  
 PS Example 3; Fig 27; 364pp; English.

XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for designing the  
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides  
 CC are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a  
 CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head  
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,  
 CC oesophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is  
 CC a savine protein of the invention  
 XX  
 XX Sequence 3541 AA;

Query Match 73.2%; Score 156; DB 5; Length 3541;  
 Best Local Similarity 100.0%; Pred. NO. 6.1e-13;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 Db 2777 LLQARLMKEESPVVSWRLEPEDGTALCFIF 2806

## RESULT 7

AAU85009

ID AAU85009 standard; peptide; 30 AA.

XX  
 AC AAU85009;

XX  
 DT 08-MAY-2002 (first entry)

XX  
 DE Human BAGE segment 1.

XX  
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 KW viral infection; human immunodeficiency virus; melanoma;  
 KW bacterial infection; Salmorella; Legionella; parasitic infection;  
 KW Trypanosoma; Toxoplasma; Giardia.

XX  
 OS Homo sapiens.

XX  
 FN WO200190197-A1.

XX  
 PD 29-NOV-2001.

XX  
 PF 25-MAY-2001; 2001WO-AU000622.

XX  
 PR 26-MAY-2000; 2000AU-00007761.

XX  
 PA (AUSU ) UNIV AUSTRALIAN NAT.

XX  
 PI Thomson SA, Ramshaw IA;

XX  
 DR WPI; 2002-147575/19.

XX  
 DR N-PSDB; ABK36829.

XX  
 PT New synthetic polypeptides having several different segments of at least  
 PT one parent polypeptide linked together differently compared to the  
 PT linkage in the parent polypeptide, for inducing immune response against a  
 PT pathogen or cancer.

XX  
 PS Example 3; Fig 27; 364pp; English.

XX  
 CC The invention relates to a new synthetic polypeptide (I) comprising  
 CC several different segments of at least one parent polypeptide linked  
 CC together in a different relationship relative to their linkage in the  
 CC parent polypeptide to impede, abrogate or otherwise alter at least one  
 CC function associated with the parent polypeptide and for inducing an  
 CC immune response against a pathogen or cancer. Also included are a  
 CC synthetic polynucleotide encoding and a computer system for designing the  
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides  
 CC are referred to as a Savine. The synthetic polypeptide is useful for



CC modulating immune responses preferably directed against a pathogen or a  
 CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head  
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,  
 CC esophagus, brain, testicle, uterus), as potentiating agents.  
 CC Compositions comprising the polypeptide may be used in the treatment or  
 CC prophylaxis against viral (such as infections caused by HIV (human  
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis  
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial  
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,  
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic  
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,  
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is  
 CC a peptide derived from a parent protein used to construct a vaccine of the  
 CC invention  
 XX

SQ Sequence 30 AA;

Query Match 59.2%; Score 126; DB 5; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-11;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLALSAQLLQARLMKESPVVS 28  
 DB 3 MAARAVFLALSAQLLQARLMKESPVVS 30

RESULT 8

BAR67808  
 ID AAR67808 standard; peptide; 22 AA.

XX AAR67808;

XX 25-MAR-2003 (revised)

DT 22-AUG-1995 (first entry)

XX BAGE tumor rejection antigen peptide.

DE BAGE; tumor rejection antigen precursor; diagnosis; HLA;  
 KW human leukocyte antigen MHC; major histocompatibility complex; TRAP;  
 KW cancer; melanoma.

XX Synthetic.

XX WO9500159-A1.

XX 05-JAN-1995.

XX 10-JUN-1994; 94WO-US006534.

XX 17-JUN-1993; 93US-00079110.

PR 15-FEB-1994; 94US-00196630.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Van Der Bruggen P, Boon-Fallieur T, Coullie P, Renauld J;

XX WPI; 1995-051741/07.

XX Nucleic acid coding for a tumour rejection antigen precursor - used to  
 PT develop prods. for the diagnosis and therapy of cancers, partic.  
 PT melanomas.

XX Claim 21; Page 19; 33pp; English.

XX This sequence encodes the tumor rejection antigen peptide BAGE. The  
 CC peptide may be used in the diagnosis and therapy of cancers, e.g.  
 CC melanomas. (Updated on 25-MAR-2003 to correct PN field.)  
 XX

SQ Sequence 22 AA;

Query Match 46.0%; Score 98; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLALSAQLLQARLMKE 22  
 DB 1 MAARAVFLALSAQLLQARLMKE 22

RESULT 9

AAV10634

ID AAY10634 standard; peptide; 22 AA.

XX AAY10634;

XX 12-MAY-1999 (first entry)

XX Peptide antigen SEQ ID NO:564.

XX Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;  
 KW malignant melanoma; viral disease; hepatitis; AIDS.

XX Synthetic.

XX Homo sapiens.

XX WO9902183-A2.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US014289.

XX 10-JUL-1997; 97CA-02209815.

PR 10-DEC-1997; 97US-00988320.

XX (CTLI-) CTL IMMUNOTHERAPIES CORP.

XX Kuendig TM, Simard JUL;

XX WPI; 1999-120514/10.

XX Inducing a cytotoxic T lymphocyte response - by maintaining a level of  
 PT antigen in the lymphatic system of a mammal so as to provide a sustained  
 PT CTL response, used to treat, e.g. AIDS.

XX Disclosure; Page 52; 199pp; English.

XX The present invention describes a method of inducing and/or sustaining an  
 CC immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
 CC method comprises: (a) delivering an antigen to the mammal at a level to  
 CC induce an immunological CTL response in the mammal; and (b) maintaining the  
 CC level of the antigen in the mammal's lymphatic system to maintain the  
 CC immunologic CTL response. The method can be used for the delivery of e.g.  
 CC a differentiation antigen, a tumour-specific multilineage antigen, an  
 CC embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene  
 CC antigen, or a viral antigen. They can be used for the treatment of  
 CC disease such as cancer, e.g. malignant melanoma or infectious disease,  
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
 CC to the lymphatic system provides for potent CTL stimulation that takes  
 CC place in the milieu of the lymphoid organ, and it sustains stimulation  
 CC that is necessary to keep CTL active, cytotoxic and recirculating through  
 CC the body. AAV10071 to AAY10639 represent examples of peptide antigens  
 CC given in the present invention  
 XX

SQ Sequence 22 AA;

Query Match 46.0%; Score 98; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLALSAQLLQARLMKE 22

DB 1 MAARAVFLALSAQLLQARLMKE 22

RESULT 10

```

ABG80319
ID ABG80319 standard; peptide; 22 AA.
XX
AC ABG80319;
XX
29-AUG-2003 (revised)
DT 15-NOV-2002 (first entry)
XX
DE MHC class I molecule, viral epitope #567.
XX
KW Major histocompatibility complex; MHC; MHC class I molecule; virus;
KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;
KW antigen; immunogenic; malignant tumor; carcinoma; melanoma; leukaemia;
KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;
KW acquired immune deficiency syndrome; AIDS.
XX
OS Viruses.
XX
FN WO200262368-A2.
XX
PD 15-AUG-2002.
XX
PF 22-JAN-2002; 2002WO-US002033.
XX
PR 02-FEB-2001; 2001US-00776232.
XX
PA (CTL1-) CTL IMMUNOTHERAPIES CORP.
XX
PI Kundig TM, Simard JLL;
XX
WPI; 2002-657506/70.
XX
Inducing or sustaining immunological cytotoxic T lymphocyte response in a
mammal, useful for treating a mammal with malignant tumor or infectious
disease, by directly administering an antigen to the lymphatic system of
the mammal.
XX
PS Disclosure; Page 46; 73pp; English.
XX
The invention relates to a method of inducing and/or sustaining an
immunological cytotoxic T lymphocyte (CTL) response in a mammal
comprising administering directly to the lymphatic system of the mammal:
(a) an antigen in the form of a polypeptide; (b) a vector comprising a
nucleic acid encoding the antigen; or (c) a non-peptide antigen. The
method is useful for inducing and/or sustaining CTL response in a mammal.
This is particularly useful for treating a mammal having a malignant
tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious
disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS),
malaria, measles or tuberculosis), or in an animal having a
disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS),
malaria, measles or tuberculosis). The mammal may be dogs, cats, mice,
cattle, sheep, pigs, goats, rabbits, or preferably humans. ABG79753-
ABG80319 represent viral epitopes on major histocompatibility complex
(MHC) class I molecules, used in the method of the invention. (Updated on
29-AUG-2003 to standardise OS field)
XX
SQ Sequence 22 AA;
Query Match 46.0%; Score 98; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.2e-07;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAARAVFLAQAQLQARLMKE 22
DB 1 MAARAVFLAQAQLQARLMKE 22
RESULT 11
AAU85011
ID AAU85011 standard; peptide; 17 AA.
XX
AC AAU85011;
XX
DT 08-MAY-2002 (first entry)

```

```

XX DE Human BAGE segment 3.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
WPI; 2002-147575/19.
XX
N-PSDB; ABK36831.
XX
New synthetic polypeptides having several different segments of at least
one parent polypeptide linked together differently compared to the
linkage in the parent polypeptide, for inducing immune response against a
pathogen or cancer.
XX
Example 3; Fig 27; 364pp; English.
XX
The invention relates to a new synthetic polypeptide (I) comprising
several different segments of at least one parent polypeptide linked
together in a different relationship relative to their linkage in the
parent polypeptide to impede, abrogate or otherwise alter at least one
function associated with the parent polypeptide and for inducing an
immune response against a pathogen or cancer. Also included are a
synthetic polynucleotide encoding and a computer system for designing the
synthetic polypeptides. The synthetic polypeptides and polynucleotides
are referred to as a Savine. The synthetic polypeptide is useful for
modulating immune responses preferably directed against a pathogen or a
cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
oesophagus, brain, testicle, uterus), as potentiating agents.
XX
Compositions comprising the polypeptide may be used in the treatment or
prophylaxis against viral (such as infections caused by HIV (human
immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
(e.g., infections caused by Neisseria, Meningococcus, Haemophilus,
Salmonella, Streptococcus, Legionella and Mycobacterium) or parasitic
(e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
a peptide derived from a parent protein used to construct a savine of the
invention
XX
SQ Sequence 17 AA;
Query Match 40.8%; Score 87; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 WLEPEDGTALCFIF 43
DB 1 WLEPEDGTALCFIF 15
RESULT 12
AAR67809
ID AAR67809 standard; peptide; 16 AA.
XX
AC AAR67809;
XX

```

DT	25-MAR-2003 (revised)
DD	22-AUG-1995 (first entry)
XX	BAGE tumor rejection antigen peptide.
XX	BAGE; tumor rejection antigen precursor; diagnosis; HLA;
KW	human leukocyte antigen YMC; major histocompatibility complex; TRAP;
KW	cancer; melanoma.
XX	Synthetic.
OS	WO9500159-A1.
XX	05-JAN-1995.
XX	10-JUN-1994; 94WO-US006534.
PF	17-JUN-1993; 93US-00079110.
PR	15-FEB-1994; 94US-00196630.
XX	(LUDW-) LUDWIG INST CANCER RES.
PA	Van Der Bruggen P, Boon-Falleur T, Coulie P, Renauld J;
XX	WPI; 1995-051741/07.
DR	Nucleic acid coding for a tumour rejection antigen precursor - used to
PT	develop prods. for the diagnosis and therapy of cancers, partic.
PT	melanomas.
XX	Claim 21; Page 20; 33pp; English.
PS	This sequence encodes the tumor rejection antigen peptide BAGE. The
XX	peptide may be used in the diagnosis and therapy of cancers, e.g.
CC	melanomas. (Updated on 25-MAR-2003 to correct PN field.)
CC	Sequence 16 AA;
XX	Query Match 32.9%; Score 70; DB 2; Length 16;
SQ	Best Local Similarity 100.0%; Pred. No. 0.0045;
	Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MAARAVFLALSQLLQ 16
D6	
	1 MAARAVFLALSQLLQ 16
	RESULT 13
AAY10635	AAY10635 standard; peptide; 16 AA.
ID	AC AAY10635;
XX	12-MAY-1999 (first entry)
DT	Peptide antigen SEQ ID NO:565.
DE	Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
XX	immunisation; tumour; infectious disease; immunotherapy; cancer;
KW	malignant melanoma; viral disease; hepatitis; AIDS.
KW	Synthetic.
OS	Homo sapiens.
XX	WO9902183-A2.
PN	21-JAN-1999.
PD	10-JUL-1998; 98WO-US014289.
PF	10-JUL-1997; 97CA-02209815.
XX	10-DEC-1997; 97US-00988320.
PR	
XX	



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:06:43 ; Search time 23 seconds  
(without alignments)  
96.518 Million cell updates/sec

Title: US-10-081-108-2  
Perfect score: 213  
Sequence: 1 MAARAVFLALSAQLQARLM.....SPVSWRLEPEDGTALCFIF 43

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	2	US-08-389-360-2
2	213	100.0	43	3	US-09-038-328-2
3	213	100.0	43	4	US-09-435-524-2
4	213	100.0	43	4	US-09-382-497-2
5	98	46.0	22	1	US-08-196-630A-8
6	70	32.9	16	1	US-08-196-630A-9
7	62	29.1	54	4	US-09-621-976-5718
8	62	29.1	293	4	US-09-205-258-328
9	62	29.1	384	4	US-09-970-516-2
10	57	26.8	398	2	US-08-284-391B-29
11	57	26.8	398	3	US-09-218-950-29
12	57	26.8	433	2	US-08-867-149-1
13	57	26.8	433	2	US-08-808-374-1
14	57	26.8	433	3	US-09-100-409A-1
15	57	26.8	458	3	US-08-466-368-4
16	57	26.8	462	2	US-08-417-495-5
17	57	26.8	462	2	US-08-284-391B-5
18	57	26.8	462	3	US-09-218-950-5
19	57	26.8	462	5	PCT-US92-01785-5
20	57	26.8	462	5	PCT-US95-00454-5
21	57	26.8	532	2	US-08-417-495-6
22	57	26.8	532	2	US-08-284-391B-6
23	57	26.8	532	3	US-09-218-950-6
24	57	26.8	532	5	PCT-US92-01785-6
25	57	26.8	532	5	PCT-US95-00454-6
26	57	26.8	575	2	US-08-417-495-4
27	57	26.8	575	2	US-08-284-391B-4

28 57 26.8 575 3 US-09-218-950-4 Sequence 4, Appli  
29 57 26.8 575 5 PCT-US92-01785-4 Sequence 4, Appli  
30 57 26.8 575 5 PCT-US95-00454-4 Sequence 4, Appli  
31 57 26.8 630 4 US-08-472-888A-6 Sequence 6, Appli  
32 56 26.3 293 4 US-09-489-039A-7884 Sequence 7, Appli  
33 54 25.4 1084 4 US-10-072-094-7 Sequence 7884, Ap  
34 54 25.4 1084 4 US-09-637-145-3 Sequence 3, Appli  
35 53.5 25.1 504 4 US-09-252-991A-29506 Sequence 29506, A  
36 53 24.9 138 4 US-09-252-991A-27704 Sequence 27704, A  
37 52.5 24.6 436 4 US-09-134-000C-3950 Sequence 3950, Ap  
38 52 24.4 87 4 US-09-489-039A-12313 Sequence 12313, A  
39 51 23.9 227 4 US-09-543-681A-6120 Sequence 6120, Ap  
40 51 23.9 404 4 US-09-543-681A-7604 Sequence 7604, Ap  
41 51 23.9 1114 4 US-09-637-145-4 Sequence 4, Appli  
42 50 23.5 1122 4 US-10-072-094-8 Sequence 8, Appli  
43 50 23.5 1122 4 US-10-072-094-91 Sequence 91, Appli  
44 49.5 23.2 247 3 US-08-619-812-2 Sequence 2, Appli  
45 49.5 23.2 355 1 US-08-471-570-14 Sequence 14, Appli

ALIGNMENTS

RESULT 1  
US-08-389-360-2  
; Sequence 2, Application US/08389360  
; Patent No. 5877017  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen et al.  
; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
; TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARES: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/389,360  
; FILING DATE: Herewith  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/196,630  
; FILING DATE: February 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/079,110  
; FILING DATE: June 17, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pasqualini Patricia A.  
; REGISTRATION NUMBER: 34,894  
; REFERENCE/DOCKET NUMBER: LUD 5310.2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-389-360-2

Query Match 100.0%; Score 213; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e-26;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 Db 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43

## RESULT 2

US-09-038-328-2  
 ; Sequence 2, Application US/09038328  
 ; Patent No. 6110694  
 ; GENERAL INFORMATION:  
 ; APPLICANT: van der Bruggen et al.  
 ; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
 ; TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
 ; TITLE OF INVENTION: THEREOF  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/038,328  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/389,360  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/079,110  
 ; FILING DATE: June 17, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pasqualini, Patricia A.  
 ; REGISTRATION NUMBER: 34,894  
 ; REFERENCE/DOCKET NUMBER: LUD 5310.2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 43  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ;

Query Match 100.0%; Score 213; DB 3; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-26;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 Db 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 RESULT 3  
 US-09-435-524-2  
 ; Sequence 2, Application US/09435524  
 ; Patent No. 6465184  
 ; GENERAL INFORMATION:  
 ; APPLICANT: van der Bruggen et al.  
 ; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
 ; TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/382,497  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/389,360  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/079,110  
 ; FILING DATE: June 17, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pasqualini, Patricia A.  
 ; REGISTRATION NUMBER: 34,894  
 ; REFERENCE/DOCKET NUMBER: LUD 5310.2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 43  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ;

Query Match 100.0%; Score 213; DB 3; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-26;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 Db 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 RESULT 3  
 US-09-435-524-2  
 ; Sequence 2, Application US/09435524  
 ; Patent No. 6465184  
 ; GENERAL INFORMATION:  
 ; APPLICANT: van der Bruggen et al.  
 ; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
 ; TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/382,497  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/389,360  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/079,110  
 ; FILING DATE: June 17, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pasqualini, Patricia A.  
 ; REGISTRATION NUMBER: 34,894  
 ; REFERENCE/DOCKET NUMBER: LUD 5310.2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 43  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ;

ADDRESSEE: Felfe & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/435,524  
 FILING DATE: 08-No. 6465184-1999  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/038,328  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 REFERENCE/DOCKET NUMBER: LUD 5310.2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 43  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ;

Query Match 100.0%; Score 213; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-26;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43  
 Db 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLEPEDGTALCFIF 43

## RESULT 4

US-09-382-497-2  
 ; Sequence 2, Application US/09382497  
 ; Patent No. 6638512  
 ; GENERAL INFORMATION:  
 ; APPLICANT: van der Bruggen et al.  
 ; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
 ; TITLE OF INVENTION: WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/382,497  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/389,360  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/079,110  
 ; FILING DATE: June 17, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pasqualini, Patricia A.  
 ; REGISTRATION NUMBER: 34,894  
 ; REFERENCE/DOCKET NUMBER: LUD 5310.2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 43  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ;

```

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/079,110
; FILING DATE: June 17, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5310.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-382-497-2
Query Match 100.0%; Score 213; DB 4; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLALSQAQLQARLMKESPVSWRLEPEDGTALCFIF 43
Db 1 MAARAVFLALSQAQLQARLMKESPVSWRLEPEDGTALCFIF 43

RESULT 5
US-08-196-630A-8
; Sequence 8, Application US/08196630A
; Patent No. 5683886
; GENERAL INFORMATION:
; APPLICANT: van der Bruggen, Pierre
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: ISOLATED PEPTIDES WHICH FORM
; TITLE OF INVENTION: COMPLEXES WITH MHC MOLECULE HLA-C-CLONE 10 AND USES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,630A
; FILING DATE: 15-FEB-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/079,110
; FILING DATE: 17-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5683886man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5310.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-196-630A-9
; Query Match 32.9%; Score 70; DB 1; Length 16;
; Best Local Similarity 100.0%; Pred. No. 0.00016;
; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLALSQAQLQ 16
Db 1 MAARAVFLALSQAQLQ 16

RESULT 7
US-09-621-976-5718
; Sequence 5718, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
;
Query Match 46.0%; Score 98; DB 1; Length 22;
US-08-196-630A-8
```





.. TITLE OF INVENTION: PROTECTION FROM INFECTION BY HIV PRIMARY ISOLATES



Search completed: July 7, 2004, 17:10:12  
Job time : 28 secs.

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:09:14 ; Search time 46 Seconds  
(without alignments)  
290.983 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213  
Sequence: 1 MAARAVFLALSAQLQALRM.....SPVSVWRLEPEDGTALCFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	12	US-10-296-734-825
2	213	100.0	43	14	Sequence 825, Appl
3	213	100.0	43	14	Sequence 81, Appl
4	156	73.2	30	12	US-10-081-108-2
5	156	73.2	3541	12	US-10-296-734-1214
6	126	59.2	30	12	US-10-296-734-1454
7	87	40.8	17	12	US-10-296-734-1212
8	62	29.1	293	10	US-09-933-767-328
9	62	29.1	293	12	US-10-004-860-328
10	62	29.1	293	14	US-10-023-282-328
11	62	29.1	333	15	US-10-264-237-2585
12	62	29.1	368	14	US-10-053-510-21
13	62	29.1	368	15	US-10-348-052-21
14	62	29.1	384	9	US-09-784-810A-2
15	62	29.1	384	9	US-09-970-516-2

Sequence 26, Appl  
Sequence 2, Appl  
Sequence 75144, A  
Sequence 3, Appl  
Sequence 72850, A  
Sequence 29, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 8269, Ap  
Sequence 3, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 8, Appl  
Sequence 7489, Ap  
Sequence 75963, A  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 30, Appl  
Sequence 36, Appl  
Sequence 2, Appl  
Sequence 7, Appl  
Sequence 12, Appl  
Sequence 7, Appl  
Sequence 4, Appl  
Sequence 15114, A

384 14 US-10-354-358-26  
384 16 US-10-619-344-2  
552 12 US-10-282-122A-75144  
384 9 US-09-796-487-3  
340 12 US-10-282-122A-72850  
398 10 US-09-939-537-29  
462 10 US-09-939-537-5  
462 11 US-09-243-008-5  
532 10 US-09-939-537-6  
532 11 US-09-243-008-6  
575 10 US-09-939-537-4  
575 11 US-09-243-008-4  
317 14 US-10-156-761-8269  
353 9 US-09-895-943-3  
353 9 US-09-895-943-3  
370 9 US-09-895-943-2  
370 9 US-09-895-943-2  
73 10 US-09-800-187-8  
336 14 US-10-032-585-7489  
552 12 US-10-282-122A-75963  
967 9 US-09-817-913-7  
967 9 US-09-817-538-7  
967 10 US-09-563-728A-30  
1030 15 US-10-115-482-36  
1084 10 US-09-800-187-2  
1084 14 US-10-072-094-7  
1084 14 US-10-173-539-12  
1084 14 US-10-172-094-7  
1084 16 US-10-360-534-4  
51 12 US-10-424-599-15114  
53.5 25.1

#### ALIGNMENTS

RESULT 1  
US-10-296-734-825  
; Sequence 825, Application US/10296734  
; Publication No. US20040054137A1  
; GENERAL INFORMATION:  
; APPLICANT: Thompson, Scott A  
; APPLICANT: Ramshaw, Ian A  
; TITLE OF INVENTION: Synthetic molecules and uses therefor  
; FILE REFERENCE: Savine  
; CURRENT APPLICATION NUMBER: US/10/296,734  
; CURRENT FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: AU PQ7761/00  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 1507  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 825  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: BAGE consensus polypeptide  
US-10-296-734-825

Query Match 100.0%; Score 213; DB 12; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.7e-23;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLALSAQLQALRMKEESPVSWSRLEPEDGTALCFIF 43  
|||||  
DB 1 MAARAVFLALSAQLQALRMKEESPVSWSRLEPEDGTALCFIF 43  
|||||

RESULT 2  
US-10-207-655-81  
; Sequence 81, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.

APPLICANT: Hayden-Ledbetter, Martha S.  
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
FILE REFERENCE: 390069.401C1  
CURRENT APPLICATION NUMBER: US/10/207,655  
CURRENT FILING DATE: 2002-07-25  
NUMBER OF SEQ ID NOS: 426  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 81  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-207-655-81

Query Match 100.0%; Score 213; DB 14; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.7e-23;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLPEPDGTALCFIF 43  
Db 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLPEPDGTALCFIF 43

RESULT 3  
US-10-081-108-2  
Sequence 2, Application US/10081108  
Publication No. US20030138854A1  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen et al.  
TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES  
WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felife & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/081,108

FILING DATE: 20-Feb-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/435,524

FILING DATE: 08-No. US20030138854A1-1999

APPLICATION NUMBER: 09/038,328

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/079,110

FILING DATE: June 17, 1993

ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.

REGISTRATION NUMBER: 34,894

REFERENCE/DOCKET NUMBER: LUD 5310.2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 43

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-081-108-2

Query Match 100.0%; Score 213; DB 14; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.7e-23;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLPEPDGTALCFIF 43  
Db 1 MAARAVFLALSAQLLQARLMKEESPVVSWRLPEPDGTALCFIF 43

RESULT 4  
US-10-296-734-1214  
Sequence 1214, Application US/10296734  
Publication No. US20040054137A1  
GENERAL INFORMATION:  
APPLICANT: Thompson, Scott A  
APPLICANT: Ramshaw, Ian A  
TITLE OF INVENTION: Synthetic molecules and uses therefor  
FILE REFERENCE: Savine  
CURRENT APPLICATION NUMBER: US/10/296,734  
CURRENT FILING DATE: 2003-08-04  
PRIOR APPLICATION NUMBER: AU PQ7761/00  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 1507  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1214  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: BAGE segment 2  
US-10-296-734-1214

Query Match 73.2%; Score 156; DB 12; Length 30;  
Best Local Similarity 100.0%; Pred. No. 6.3e-15;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKEESPVVSWRLPEPDGTALCFIF 43  
Db 1 LLQARLMKEESPVVSWRLPEPDGTALCFIF 30

RESULT 5  
US-10-296-734-1454  
Sequence 1454, Application US/10296734  
Publication No. US20040054137A1  
GENERAL INFORMATION:  
APPLICANT: Thompson, Scott A  
APPLICANT: Ramshaw, Ian A  
TITLE OF INVENTION: Synthetic molecules and uses therefor  
FILE REFERENCE: Savine  
CURRENT APPLICATION NUMBER: US/10/296,734  
CURRENT FILING DATE: 2003-08-04  
PRIOR APPLICATION NUMBER: AU PQ7761/00  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 1507  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1454  
LENGTH: 3541  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Melanoma cancer specific savine  
US-10-296-734-1454

Query Match 73.2%; Score 156; DB 12; Length 3541;  
Best Local Similarity 100.0%; Pred. No. 1.7e-12;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKEESPVVSWRLPEPDGTALCFIF 43  
Db 2777 LLQARLMKEESPVVSWRLPEPDGTALCFIF 2806

RESULT 6  
US-10-296-734-1212

```
; Sequence 1212, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1212
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: BAGE segment 1
; US-10-296-734-1212

Query Match          59.2%; Score 126; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQAQLQARLMKEESPVVS 28
Db 3 MAARAVFLAQAQLQARLMKEESPVVS 30

RESULT 7
US-10-296-734-1216
; Sequence 1216, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU P07761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1216
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: BAGE segment 3
; US-10-296-734-1216

Query Match          40.8%; Score 87; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 WRLEPEDGTALCFIF 43
Db 1 WRLEPEDGTALCFIF 15

RESULT 8
US-09-933-767-328
; Sequence 328, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
```









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 17:05:13 ; Search time 20 Seconds  
(without alignments)  
206.812 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 YARAVFLALSAQLQARLM.....SPVSVWRLEPEDGTALCFIP 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.5	26.1	359	2 JC7280	cytokine receptor-
2	55	25.8	154	2 F69520	conserved hypot
3	54.5	25.6	994	1 TQECN5	transposase, Esch
4	54	25.4	295	2 G80934	probable exinucle
5	54	25.4	295	2 C85783	hypothetical prote
6	54	25.4	295	2 E64933	conserved hypotet
7	54	25.4	302	2 AD0709	conserved hypotet
8	54	25.4	525	2 T00459	hypothetical prote
9	54	25.4	552	2 AC0609	conserved hypotet
10	54	25.4	699	2 A96802	unknown protein [i
11	53.5	25.1	155	2 E97591	structure homolog
12	53	24.9	552	2 D64826	ybjd protein - Esc
13	53	24.9	552	2 B90749	hypothetical prote
14	53	24.9	552	2 F85599	hypothetical prote
15	52.5	24.6	243	2 T04489	protein kinase hom
16	52	24.4	346	2 T51377	hypothetical prote
17	52	24.4	445	2 T05639	hypothetical prote
18	52	24.4	945	2 A64714	helicase - Helicob
19	51	23.9	254	2 T09009	cd3-like protein
20	51	23.9	284	2 G70732	probable thiosulfa
21	51	23.9	332	2 A99738	probable membrane
22	51	23.9	332	2 B85588	hypothetical prote
23	51	23.9	332	2 C64816	hypothetical prote
24	51	23.9	524	1 S36175	glycerol kinase (E
25	51	23.9	524	2 JNC606	ATP-stimulated glu
26	51	23.9	946	2 A71805	probable ATP-depen
27	50.5	23.7	862	2 AC1214	fibrinogen-binding
28	50	23.5	58	2 S13979	hypothetical prote
29	50	23.5	270	2 A72800	gpi protein - Myco

30 50 23.5 334 2 G75344 probable polyferre  
31 50 23.5 367 2 AI2606 hypothetical prote  
32 50 23.5 367 2 H97388 periplasmic mannit  
33 50 23.5 396 2 AH2275 two-component resp  
34 50 23.5 503 2 T48825 hypothetical prote  
35 49.5 23.2 151 2 AB0331 sigma E factor reg  
36 49.5 23.2 247 2 A43846 lipoprotein lppA -  
37 49.5 23.2 254 2 C42691 fibroblast growth  
38 49.5 23.2 769 2 S16236 fibroblast growth  
39 49.5 23.2 821 1 TVHUF2 fibroblast growth  
40 49.5 23.2 822 2 A45081 fibroblast growth  
41 49.5 23.2 822 2 A41794 keratinocyte growt  
42 49 23.0 267 2 S74899 extragenic suppres  
43 49 23.0 319 2 T27862 hypothetical prote  
44 49 23.0 353 2 B72468 hypothetical prote  
45 49 23.0 411 2 B83782 flavohemoglobin hm

#### ALIGNMENTS

##### RESULT 1

JC7280

cytokine receptor-like molecule-2 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C;Accession: JC7280

R;Hiroyama, T.; Iwama, A.; Morita, Y.; Nakamura, Y.; Shibuya, A.; Nakauchi, H.

Biochem. Biophys. Res. Commun. 272, 224-229, 2000

A;Title: Molecular cloning and characterization of CRLM-2, a novel type I cytokine rece

A;Reference number: JC7280

A;Contents: Embryo

A;Accession: JC7280

A;Molecule type: mRNA

A;Residues: 1-359 <HIR>

A;Cross-references: DDBJ:AB039945

C;Genetics:

A;Gene: crlm-2

C;Keywords: cytokine; embryo; receptor; signal transduction; transmembrane protein

Query Match

Best Local Similarity 26.1%; Score 55.5; DB 2; Length 359;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

QY 21 KBESPW---SWRLPEPDGTALCFI 42

Db 301 EREDDLIHTKAKRVEPDGTSUCTV 325

##### RESULT 2

F69520

conserved hypothetical protein AF2166 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C;Accession: F69520

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: F69520

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-154 <KLE>

A;Cross-references: GB:AE0000955; GB:AE000782; NID:G689278; PID:AAB89089.1; PID:G26483

Query Match

Best Local Similarity 25.8%; Score 55; DB 2; Length 154;

Matches 13; Conservative 10; Mismatches 15; Indels 0; Gaps 0;



C;Genetics:  
A;Gene: STY1804

Query Match 25.4%; Score 54; DB 2; Length 302;  
Best Local Similarity 25.0%; Pred. No. 13;  
Matches 11; Conservative 5; Mismatches 0; Gaps 0;

QY 10 LSAQLIQARLMKEESPVSWSRL 31  
: |||:||||:|:|  
DB 98 MGALLLEARLIKEQFLFNKRL 119

RESULT 8  
T00459

hypothetical protein T14N5.16 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana [mouse-ear cress]  
C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 22-Oct-1999  
C;Accession: T00459  
R;Fedeerspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Araujo, J.; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.  
submitted to the EMBL Data Library, September 1998  
A;Reference number: Z14152  
A;Accession: T00459  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-525 <FED>  
A;Cross-references: EMBL:AC004260; NID:g3176694; GSPDB:GN00059; ATSP:T14N5  
A;Experimental source: cultivar Columbia  
C;Genetics:  
A;Gene: ATSP:T14N5.16  
A;Map position: 1  
A;Introns: 83/3; 112/3; 143/3; 169/2; 212/3; 335/3; 363/1; 400/3; 429/3; 471/3

Query Match 25.4%; Score 54; DB 2; Length 525;  
Best Local Similarity 40.0%; Pred. No. 23;  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 6 VFLALSALQLIQAALRMKEESPWSWR 30  
: |||:||||:|:|  
DB 25 ILTAGDRKLLKVELLKEETLVSWK 49

RESULT 9  
AC0609

conserved hypothetical protein STY0935 [imported] - Salmonella enterica subsp. enterica  
C;Species: Salmonella enterica subsp. enterica serovar Typhi  
A;Note: this species has also been called Salmonella typhi  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AC0609  
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AC0609  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-552 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD05341.1; PID:g16502105; GSPDB:GN00176  
C;Genetics:  
A;Gene: STY0935

Query Match 25.4%; Score 54; DB 2; Length 552;  
Best Local Similarity 34.9%; Pred. No. 24;  
Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

QY 11 SAQLIQ-----ARLMKEESPVSWSR-----LEPDGTALCF 41  
: |||:||||:|:|  
DB 324 SGELLSLTPVHVVRVLRSSRVAAARLGTGGLSADGRIAF 366

**RESULT 10**

A96802  
unknown protein [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: A96802  
R/Theologas, A.; Ecker, J.R.; Palm, C.J.; Pedersenpiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Autheors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Autheors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Ritle: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; PMID:21016719; PMID:11130712  
A/Accession: A96802  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-699 <STO>  
A/Cross-references: GB:AE005173; NID:g11079480; PIDN:AAG29192.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: F2P24.2  
A/Map position: 1

Query Match 25.4%; Score 54; DB 2; Length 699;  
Best Local Similarity 40.0%; Pred.No. 31;  
Matches 10; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

**QY** 6 VFPLALSQAQLLQARLMKEESPVSWSR 30  
:  
  
**Db** 25 ILTPAGDRKLLKVELLXETTIVSWK 49  
:  
:::  
:::  
:::  
:::  
:::  
:::

**RESULT 11**

E97591  
structure homolog of thermus thermophilus 30S ribosomal chain M [imported] - Agrobacter  
C/Species: Agrobacterium tumefaciens  
C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 12-Jun-2003  
C/Accession: E97591  
R/Goodner, B.; Hankle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourullo, B.; Goldman  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markeiz, B.  
Science 294, 2323-2328, 2001  
A/Ritle: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A/Reference number: A97359; PMID:21608551; PMID:11743194  
A/Accession: E97591  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-155 <KUR>  
A/Cross-references: GB:AE007869; PIDN:AAK87686.1; PID:g15157045; GSPDB:GN00169  
C/Genetics:  
A/Gene: AGR\_C\_3520  
A/Map position: circular chromosome  
C/Superfamily: ribosomal protein Sl3/S18

Query Match 25.1%; Score 53.5; DB 2; Length 155;  
Best Local Similarity 54.5%; Pred.No. 7.4;  
Matches 12; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

**QY** 12 AQLLQARLMKEESPVSWSRLRP 33  
:  
  
**Db** 11 AVLLAARM---ESPLAGWQLEP 29  
:  

**RESULT 12**

D64826  
yjd protein - Escherichia coli (strain K-12)  
C/Species: Escherichia coli  
C/Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C/Accession: D64826  
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C  
A.; Rose, D.J.; Mau, B.; Shao, Y.

100

Science 277, 1453-1462, 1997  
A>Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: D64826  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-552 <BLAT>  
A:Cross-references: GB:AB000189; GB:U00096; NID:gl787097; PIDN:AAC73963.1; PID:gl787102;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:

A:Gene: ybJD  
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein  
C:29-36/Region: nucleotide-binding motif A (P-loop)  
F:264-280/Domain: transmembrane #status predicted <TM>

Query Match 24.9%; Score 53; DB 2; Length 552;  
Best Local Similarity 34.9%; Pred. No. 33;  
Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

QY 11 SAQLIQ-----ARLMKESPVVSRLP-----EDGTALCF 41  
DB 324 SGEILLSTPVEHVCLVRESRVAAWRLGPSGLSTEDSRISF 366

## RESULT 13

B90749  
hypothetical protein ECs0962 [imported] - *Escherichia coli* (strain O157:H7, substrain RIMD 0509952)  
C:Species: *Escherichia coli*  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: B90749  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gen-  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: B90749  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-552 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA334385.1; PID:gl33360421; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs0962

Query Match 24.9%; Score 53; DB 2; Length 552;  
Best Local Similarity 34.9%; Pred. No. 33;  
Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

QY 11 SAQLIQ-----ARLMKESPVVSRLP-----EDGTALCF 41  
DB 324 SGEILLSTPVEHVCLVRESRVAAWRLGPSGLSTEDSRISF 366

## RESULT 14

R85599  
hypothetical protein ybJD [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)  
C:Species: *Escherichia coli*  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: R85599  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: R85599  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-552 <STO>  
A:Cross-references: GB:AE005174; NID:gl2513906; PIDN:AAG55258.1; GSPDB:GN00145; UWGP:Z11  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: ybJD

Query Match 24.9%; Score 53; DB 2; Length 552;  
Best Local Similarity 34.9%; Pred. No. 33;  
Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

QY 11 SAQLIQ-----ARLMKESPVVSRLP-----EDGTALCF 41  
DB 324 SGEILLSTPVEHVCLVRESRVAAWRLGPSGLSTEDSRISF 366

## RESULT 15

T04489  
protein kinase homolog F8F16.40 - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999  
C:Accession: T04489  
R:Sevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Mewe  
submitted to the Protein Sequence Database, April 1998  
A:Reference number: Z15375  
A:Accession: T04489  
A:Molecule type: DNA  
A:Residues: 1-243 <BEV>  
A:Cross-references: ENBL:AL021633  
A:Experimental source: cultivar Columbia; BAC clone F8F16  
C:Genetics:  
A:Map position: 4  
A:Note: F8F16.40

Query Match 24.6%; Score 52.5; DB 2; Length 243;  
Best Local Similarity 39.4%; Pred. No. 16;  
Matches 13; Conservative 7; Mismatches 8; Indels 5; Gaps 2;

QY 15 LQARLMKE-ESPVSWSW-----RLEPEDGTALCFI 42  
DB 24 LEDRLFRQGDSPALSWQTRFRIAAEIGTVLFL 56

Search completed: July 7, 2004, 17:09:33  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 16:58:33 ; Search time 12 Seconds  
(without alignments)  
186.585 Million cell updates/sec

Title: US-10-081-108-2  
Perfect score: 213  
Sequence: 1 MAARAVFLAALQQLQRLM.....SPVSVWRLEPDEOTALCFIF 43

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	1 BGE1_HUMAN	Q13072 homo sapien
2	206	96.7	43	1 BGE5_HUMAN	Q86Y27 homo sapien
3	177	83.1	109	1 BGE2_HUMAN	Q86Y30 homo sapien
4	177	83.1	109	1 BGE3_HUMAN	Q86Y29 homo sapien
5	173	81.2	39	1 BGE4_HUMAN	Q86Y28 homo sapien
6	62	29.1	384	1 SPH1_HUMAN	Q9NV41 homo sapien
7	55.5	26.1	359	1 CRL2_MOUSE	Q8CI19 mus musculu
8	55	25.8	154	1 YL66_ARCFU	Q28116 archaeoglob
9	54.5	25.6	994	1 TNP5_ECOLI	P08504 escherichia
10	54	25.4	293	1 CHO_SALTY	Q8Z6G5 salmonella
11	54	25.4	293	1 CHO_SALTY	Q8Z6G5 salmonella
12	54	25.4	295	1 CHO_ECO57	Q8XDX7 escherichia
13	54	25.4	295	1 CHO_ECOL6	P59439 escherichia
14	54	25.4	295	1 CHO_ECOLI	P76213 escherichia
15	54	25.4	295	1 CHO_SHIFL	P59361 shigella fl
16	54	25.4	1080	1 HDA4_CHICK	P83038 gallus gall
17	54	25.4	1084	1 HDA4_HUMAN	P56524 homo sapien
18	53	24.9	552	1 YBJD_ECOLI	P75828 escherichia
19	52.5	24.6	170	1 YC66_MESVI	Q9MUC8 mesostigma
20	52	24.4	332	1 YHGG_ECOL6	Q8FJ56 escherichia
21	51	23.9	284	1 THT3_MYCTU	Q59570 mycobacteri
22	51	23.9	332	1 YHGG_ECO57	Q8X7Y9 escherichia
23	51	23.9	332	1 YHGG_ECOLI	P75777 escherichia
24	51	23.9	332	1 YHGG_SHIFL	Q83S36 shigella fl
25	51	23.9	457	1 CD4_SALISC	Q29037 saimiri sci
26	51	23.9	524	1 GLPK_HUMAN	P32189 homo sapien
27	51	23.9	524	1 GLPK_MOUSE	Q84516 mus musculu
28	51	23.9	524	1 GLPK_RAT	Q83060 rattus norv
29	51	23.9	1113	1 HDA5_MOUSE	Q92ZV6 mus musculu
30	50	23.5	270	1 VG01_BPMD2	Q64197 mycobacteri
31	50	23.5	708	1 TRAB_HUMAN	Q9UG10 homo sapien
32	50	23.5	1122	1 HDA5_HUMAN	Q9UG16 homo sapien
33	49.5	23.2	699	1 M1B1_HUMAN	Q9UKM7 homo sapien

## ALIGNMENTS

## RESULT 1

ID	BGE1_HUMAN	STANDARD;	PRT;	43 AA.
AC	Q13072;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	B melanoma antigen 1 precursor (B melanoma antigen) (Antigen M22-BA).			
GN	BAGE OR BAGE1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM BAGE1A).			
RC	TISSUE=Melanoma;			
RX	MEDLINE=955202592; PubMed=7895173;			
RA	Boel P., Wildmann C., Sensi M.L., Brasseur R., Renaud J.-C.,			
RA	Coulie P., Boon T., van der Bruggen P.;			
RT	"BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes.";			
RL	Immunity 2:167-175(1995).			
RN	[2]			
RP	ALTERNATIVE SPLICING.			
RC	MEDLINE=22349465; PubMed=12461691;			
RA	Ruault M., van der Bruggen P., Brun M.-E., Boyle S., Roizes G.,			
RA	De Sario A.;			
RT	"New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile.";			
RL	Eur. J. Hum. Genet. 10:833-840(2002).			
CC	-!- FUNCTION: Unknown. Antigen recognized on a melanoma by autologous cytolytic T lymphocytes.			
CC	-!- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=1;			
CC	Comment=At least 6 different mRNAs, BAGE1a, BAGE1b, BAGE1c, BAGE1d, BAGE1e and BAGE1f, are produced by alternative splicing. They differ in the 3' region, but give rise to the same protein sequence;			
CC	Name=1; Synonyms=BAGE1a;			
CC	Isoid=Q13072-1; Sequence=Displayed;			
CC	-!- TISSUE SPECIFICITY: Not expressed in normal tissues, except in testis. Expressed with significant proportion in melanomas, but also in tumors of various histological origins, such as bladder carcinomas, head and neck squamous cell carcinomas, lung and breast carcinomas. Not expressed in renal, colorectal and prostatic carcinomas, leukemias and lymphomas. More frequently expressed in metastatic melanomas than in primary melanomas.			
CC	-!- MISCELLANEOUS: The ancestral BAGE gene was generated by juxtacentromeric reshuffling of the MLL3 gene. The BAGE family was expanded by juxtacentromeric movement and/or acrocentric exchanges. BAGE family is composed of expressed genes that map to the juxtacentromeric regions of chromosomes 13 and 21 and of unexpressed gene fragments that scattered in the juxtacentromeric regions of several chromosomes, including chromosomes 9, 13, 18 and 21.			

P21802 homo sapien  
Q32861 mycoplasma  
P01730 homo sapien  
P16004 pan troglod  
Q94738 strongyloce  
Q06068 strongyloce  
P05703 escherichia  
Q52042 escherichia  
Q52043 escherichia  
P31471 escherichia  
P20584 aspergillus  
O14238 schizosacch

34 49.5 23.2 821 1 FGR2\_HUMAN  
35 49 23.0 351 1 FTSY\_MYCHO  
36 49 23.0 458 1 CD4\_HUMAN  
37 49 23.0 458 1 CD4\_PANTR  
38 49 23.0 886 1 HS97\_STRPN  
39 49 23.0 889 1 HS97\_STRPU  
40 48.5 22.8 101 1 CCDB\_ECOLI  
41 48.5 22.8 126 1 CCB3\_ECOLI  
42 48.5 22.8 126 1 CCB4\_ECOLI  
43 48.5 22.8 400 1 YIEL\_ECOLI  
44 48.5 22.8 436 1 PPA1\_ASPNG  
45 48 22.5 122 1 OSTE\_SCHPO

CC -!- SIMILARITY: Belongs to the BAGE family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U19180; AAC50123.1; -.  
 CC EMBL; AF499647; AAP30743.1; -.  
 CC EMBL; AF527550; AAO32670.1; -.  
 CC EMBL; AF527551; AAO32671.1; -.  
 CC EMBL; AF527552; AAO32672.1; -.  
 CC EMBL; AF527553; AAO32673.1; -.  
 CC EMBL; AF527554; AAO32674.1; -.  
 CC Genew; HGNC:942; BAGE.  
 CC MIM; 605167; -.  
 CC Antigen; Alternative splicing; Multigene family; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 43 B MELANOMA ANTIGEN 1.  
 CC SEQUENCE 43 AA; 4810 MW; 36F3BBCE4012F1BB CRC64;  
 CC  
 CC Query Match 100.0%; Score 213; DB 1; Length 43;  
 CC Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
 CC Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MAARAVFLALSAQLLQARLMKEESPVSVWRLEPEDGTALCFIF 43  
 CC DB 1 MAARAVFLALSAQLLQARLMKEESPVSVWRLEPEDGTALCFIF 43

RESULT 2  
 BGES HUMAN  
 ID BGES HUMAN STANDARD; PRT; 43 AA.  
 AC Q86V27;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE B melanoma antigen 5 precursor.  
 GN BAGE5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=22349465; PubMed=12461691;  
 RA Ruault M., Van Der Bruggen P., Brun M.-E., Boyle S., Roizes G.,  
 RA Sario A.D.;  
 RT "New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric  
 RT regions of human chromosomes 13 and 21 have a cancer/testis expression  
 RT profile.";  
 RL Eur. J. Hum. Genet. 10:833-840(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=22563396; PubMed=12676563;  
 RA Ruault M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,  
 RA Roizes G., De Sario A.;  
 RT "BAGE genes generated by juxtacentromeric reshuffling in the hominidae  
 RT lineage are under selective pressure.";  
 RL Genomics 81:391-399(2003).  
 CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens.  
 CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in  
 CC testis. Expressed in melanoma, bladder and lung carcinomas.  
 CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by  
 CC juxtacentromeric reshuffling of the MLI3 gene. The BAGE family was  
 CC expanded by juxtacentromeric movement and/or acrocentric  
 CC exchanges. BAGE family is composed of expressed genes that map to  
 CC the juxtacentromeric regions of chromosomes 13 and 21 and of

CC unexpressed gene fragments that scattered in the juxtacentromeric  
 CC regions of several chromosomes, including chromosomes 9, 13, 18  
 CC and 21.  
 CC -!- SIMILARITY: Belongs to the BAGE family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF339516; AAO32636.1; -.  
 CC Genew; HGNC:15732; BAGE5.  
 CC Antigen; Multigene family; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 43 B MELANOMA ANTIGEN 5.  
 CC SEQUENCE 43 AA; 4710 MW; C6F3AAACE4012F1A6 CRC64;  
 CC  
 CC Query Match 96.7%; Score 206; DB 1; Length 43;  
 CC Best Local Similarity 97.7%; Pred. No. 1e-21;  
 CC Matches 42; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MAARAVFLALSAQLLQARLMKEESPVSVWRLEPEDGTALCFIF 43  
 CC DB 1 MAAGAVFLALSAQLLQARLMKEESPVSVWRLEPEDGTALCFIF 43

RESULT 3  
 BGES HUMAN  
 ID BGES HUMAN STANDARD; PRT; 109 AA.  
 AC Q86Y30;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE B melanoma antigen 2 precursor.  
 GN BAGE2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=22349465; PubMed=12461691;  
 RA Ruault M., Van Der Bruggen P., Brun M.-E., Boyle S., Roizes G.,  
 RA Sario A.D.;  
 RT "New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric  
 RT regions of human chromosomes 13 and 21 have a cancer/testis expression  
 RT profile.";  
 RL Eur. J. Hum. Genet. 10:833-840(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=22563396; PubMed=12676563;  
 RA Ruault M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,  
 RA Roizes G., De Sario A.;  
 RT "BAGE genes generated by juxtacentromeric reshuffling in the hominidae  
 RT lineage are under selective pressure.";  
 RL Genomics 81:391-399(2003).  
 CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens.  
 CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in  
 CC testis. Expressed in 22% of melanomas, in bladder and lung  
 CC carcinomas.  
 CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by  
 CC juxtacentromeric reshuffling of the MLI3 gene. The BAGE family was  
 CC expanded by juxtacentromeric movement and/or acrocentric  
 CC exchanges. BAGE family is composed of expressed genes that map to  
 CC the juxtacentromeric regions of chromosomes 13 and 21 and of  
 CC unexpressed gene fragments that scattered in the juxtacentromeric  
 CC regions of several chromosomes, including chromosomes 9, 13, 18  
 CC and 21.



CC -!- SIMILARITY: Belongs to the BAGE family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF218570; AAU55648.1; -.  
 CC Genew; HGNC:15723; BAGE2.  
 CC Antigen; Multigene family; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 109 B MELANOMA ANTIGEN 2.  
 CC SEQUENCE 109 AA; 12114 MW; 51737AR2C2135FF CRC64;  
 CC  
 CC Query Match 83.1%; Score 177; DB 1; Length 109;  
 CC Best Local Similarity 94.9%; Pred. No. 2.7e-17;  
 CC Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MAARAVFLALSQAQLQARLMKEESPVSVRLEPEDGTAL 39  
 CC DB 1 MAAGVFLALSQAQLQARLMKEESPVSVRLEPEDGTAL 39  
 CC  
 CC RESULT 4  
 CC BGE3 HUMAN STANDARD; PRT; 109 AA.  
 CC ID\_BGE3\_HUMAN  
 CC AC Q86Y29;  
 CC DT 10-OCT-2003 (Rel. 42, Created)  
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 CC DE B melanoma antigen 2 precursor.  
 CC GN BAGE3.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=22349465; PubMed=12461691;  
 CC RA Ruault M., Van Der Bruggen P., Brun M.-E., Roizes G.,  
 CC Sario A.D.;  
 CC "New BAGE (B melanoma antigen) genes mapping to the juxtaacentromeric  
 CC regions of human chromosomes 13 and 21 have a cancer/testis expression  
 CC profile.";  
 CC Eur. J. Hum. Genet. 10:833-840(2002).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=22563396; PubMed=12676563;  
 CC RA Ruault M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,  
 CC Roizes G., De Sario A.;  
 CC "BAGE genes generated by juxtaacentromeric reshuffling in the hominidae  
 CC lineage are under selective pressure.";  
 CC Genomics 81:391-399(2003).  
 CC RL  
 CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens.  
 CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in  
 CC testis. Expressed in melanoma, bladder and lung carcinomas.  
 CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by  
 CC juxtaacentromeric reshuffling of the MLL3 gene. The BAGE family was  
 CC expanded by juxtaacentromeric movement and/or acrocentric  
 CC exchanges. BAGE family is composed of expressed genes that map to  
 CC the juxtaacentromeric regions of chromosomes 13 and 21 and of  
 CC unexpressed gene fragments that scattered in the juxtaacentromeric  
 CC regions of several chromosomes, including chromosomes 9, 13, 18  
 CC and 21.  
 CC -!- SIMILARITY: Belongs to the BAGE family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF339514; AAO32634.1; -.  
 CC Genew; HGNC:15728; BAGE3.  
 CC Antigen; Multigene family; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 109 B MELANOMA ANTIGEN 3.  
 CC SEQUENCE 109 AA; 12112 MW; A477B7A8FC3A2F4E CRC64;  
 CC  
 CC Query Match 83.1%; Score 177; DB 1; Length 109;  
 CC Best Local Similarity 94.9%; Pred. No. 2.7e-17;  
 CC Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MAARAVFLALSQAQLQARLMKEESPVSVRLEPEDGTAL 39  
 CC DB 1 MAAGVFLALSQAQLQARLMKEESPVSVRLEPEDGTAL 39  
 CC  
 CC RESULT 5  
 CC BGE4 HUMAN STANDARD; PRT; 39 AA.  
 CC ID\_BGE4\_HUMAN  
 CC AC Q86Y28;  
 CC DT 10-OCT-2003 (Rel. 42, Created)  
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 CC DE B melanoma antigen 4 precursor.  
 CC GN BAGE4.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Testis;  
 CC RA Ruault M., Roichi M., Boyle S., Roizes G., van der Bruggen P.,  
 CC De Sario A.;  
 CC "BAGE a family of centromeric genes coding for tumor antigens.";  
 CC Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens.  
 CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in  
 CC testis. Expressed in melanoma, bladder and lung carcinomas.  
 CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by  
 CC juxtaacentromeric reshuffling of the MLL3 gene. The BAGE family was  
 CC expanded by juxtaacentromeric movement and/or acrocentric  
 CC exchanges. BAGE family is composed of expressed genes that map to  
 CC the juxtaacentromeric regions of chromosomes 13 and 21 and of  
 CC unexpressed gene fragments that scattered in the juxtaacentromeric  
 CC regions of several chromosomes, including chromosomes 9, 13, 18  
 CC and 21.  
 CC -!- SIMILARITY: Belongs to the BAGE family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF339515; AAO32635.1; -.  
 CC Genew; HGNC:15730; BAGE4.  
 CC Antigen; Multigene family; Signal.  
 CC SIGNAL 1 17 POTENTIAL.  
 CC CHAIN 18 39 B MELANOMA ANTIGEN 4.  
 CC SEQUENCE 39 AA; 4230 MW; 0016A0A6D0266FFF CRC64;  
 CC  
 CC Query Match 81.2%; Score 173; DB 1; Length 39;  
 CC Best Local Similarity 94.9%; Pred. No. 3.2e-17;  
 CC Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MAARAVFLALSQAQLQARLMKEESPVSVRLEPEDGTAL 39

```

Db 1 MAAGAVFLALSAQLLQARLMKEESPVSWSWLEPEDGTAL 39
|||||
RESULT 6
ID SPH1 HUMAN STANDARD; PRT; 384 AA.
AC Q9NYA1, Q9HD92; Q9NY70, Q9NYL3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sphingosine kinase 1 (EC 2.7.1.-) (SK 1) (SPK 1).
GN SPHK1 OR SPKH OR SPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20323213; PubMed=10863092;
RA Melendez A.J., Carlos-Dias E., Gosink M., Allen J.M., Takacs L.;
RT "Human sphingosine kinase: molecular cloning, functional
RT characterization and tissue distribution.";
RL Gene 251:19-26(2000).
[2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20263733; PubMed=10802064;
RA Nava V.E., Lacana E., Poulton S.;
RA Milstien S., Kohama T., Spiegel S.;
RT "Functional characterization of human sphingosine kinase-1.";
RL FEBS Lett. 473:81-84(2000).
[3]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20407120; PubMed=10947957;
RA Pitson S.M., D'Andrea R.J., Vandeleur L., Moretti P.A.B., Xia P.,
RA Gamble J.R., Vadas M.A., Wattenberg B.W.;
RT "Human sphingosine kinase: purification, molecular cloning and
RT characterization of the native and recombinant enzymes.";
RL Biochem. J. 350:429-441(2000).
[4]
RP SEQUENCE FROM N.A.
RA Van Velthoven P.P., Gijssels S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA TISSUE=Mammary gland, and Ovary;
RA Isoqai T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
RA Nagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Iehida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Iehi S., Kawai Y., Saito K.,
RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the phosphorylation of sphingosine to form
CC sphingosine 1-phosphate (SPH), a lipid mediator with both intra-
CC and extracellular functions. Also acts on D-erythro-sphingosine
CC and to a lesser extent sphinganine, but not other lipids, such as
CC D,L-threo-dihydrosphingosine, N,N-dimethylsphingosine,
CC diacylglycerol, ceramide, or phosphatidylinositol.
CC -!- CATALYTIC ACTIVITY: Sphingosine + ATP = sphingosine 1-phosphate +
CC ADP.
CC -!- SUBUNIT: Binds to calmodulin.
CC -!- TISSUE SPECIFICITY: Widely expressed with highest levels in adult
CC liver, kidney, heart and skeletal muscle.
CC -!- SIMILARITY: Contains 1 DAGKC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF266756; AAF73470.1; -
DR EMBL; AF238083; AAF73423.1; -
DR EMBL; AF200328; AAG01980.1; -
DR EMBL; AK023393; BAB14558.1; -
DR EMBL; AK022402; BAB14028.1; -
DR EMBL; AJ245504; CAB92131.1; -
DR Genew; HGNC:11240; SPK1.
DR MIM; 603730; -
DR GO; GO:0005829; C:cytosol; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0005524; P:ATP binding; IDA.
DR GO; GO:0005516; P:calmodulin binding; IDA.
DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; IDA.
DR GO; GO:0000287; F:magnesium ion binding; IDA.
DR GO; GO:0006916; P:anti-apoptosis; TAS.
DR GO; GO:0019722; P:calcium-mediated signaling; IDA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.
DR GO; GO:0007242; P:intracellular signaling cascade; TAS.
DR GO; GO:0045766; P:positive regulation of angiogenesis; IDA.
DR GO; GO:0030307; P:positive regulation of cell growth; IDA.
DR GO; GO:0030335; P:positive regulation of cell migration; IDA.
DR GO; GO:0045931; P:positive regulation of mitotic cell cycle; IDA.
DR GO; GO:0045987; P:positive regulation of smooth muscle contra. .; IDA.
DR GO; GO:0046521; P:sphingoid catabolism; NAS.
DR InterPro; IPR001206; DAGKC.
DR Pfam; PF00781; DAGKC; 1.
DR Prodom; PD005043; DAGKC; 1.
DR SMART; SM00046; DAGKC; 1.
KW Transferase; Kinase; ATP-binding; Calmodulin-binding.
FT CONFLICT 6 6 MISSING (IN REF. 4).
FT CONFLICT 11 15 LPRPC -> AXL (IN REF. 4).
FT CONFLICT 114 115 NA -> KP (IN REF. 4).
FT CONFLICT 231 251 V -> M (IN REF. 2).
FT CONFLICT 260 260 V -> I (IN REF. 2).
FT CONFLICT 302 302 L -> F (IN REF. 2).
FT CONFLICT 325 325 V -> G (IN REF. 4).
FT CONFLICT 337 337 V -> M (IN REF. 3).
SQ SEQUENCE 384 AA; 42517 MW; EB04A7F2034C2DB0 CRC64;
Query Match 29.1%; Score 62; DB 1; Length 384;
Best Local Similarity 40.0%; Pred. No. 0.64;
Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;
QY 3 ARAVFLALSAQLLQARLMKEES-----PVVSWRLEPEDG 36
DB 295 SRAMLLRLFLAMEKGRHNEYCPVLVYVVPVAFRLPEKDG 334
|||||
RESULT 7
CR12 MOUSE STANDARD; PRT; 359 AA.
ID CR12 MOUSE STANDARD; PRT; 359 AA.
AC Q8CJ19; Q9CRJ6; Q9JIE7; Q9JIO7; Q9JH8; Q9JMD5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Cytokine receptor-like factor 2 precursor (Type I cytokine receptor
DE delta 1) (Cytokine receptor-like molecule-2) (CRLM-2) (Thymic stromal
DE lymphopoietin protein receptor) (TSLPR).
GN CRL2 OR CRLM2 OR TSLPR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Embryo;
RX MEDLINE=20329232; PubMed=10872831;
RA Hiroshima T., Iwama A., Morita Y., Nakamura Y., Shibuya A.,
RA Nakauchi H.;
RT "Molecular cloning and characterization of CRLM-2, a novel type I

```

cytokine receptor preferentially expressed in hematopoietic cells.";  
[2] Biochem. Biophys. Res. Commun. 272:224-229 (2000).  
[3] SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.  
TISSUE=Lymphocytes;  
MEDLINE=20137856; PubMed=10733486;  
Fujio K., Nosaka T., Kojima T., Kawashima T., Yahata T.,  
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Yamamoto K., Nishimura T.,  
RA Kitamura T.;  
RA "Molecular cloning of a novel type I cytokine receptor similar to the  
RA common gamma chain.";  
RA Blood 95:2204-2211 (2000).  
[4] SEQUENCE FROM N.A. (ISOFORM 1).  
STRAIN=C57BL/6; TISSUE=Lymphocytes;  
RA MEDLINE=20432254; PubMed=10574032;  
RA Park L.S., Martin U., Garka K., Gliniak B., Di Santo J.P., Muller W.,  
RA Langespade D.A., Copeland N.G., Jenkins N.A., Farr A.G.,  
RA Ziegler S.F., Morrissey P.J., Paxton R., Sims J.E.;  
RA "Cloning of the murine thymic stromal lymphopoietin (TSLP) receptor.  
RA Formation Of a functional heteromeric complex requires interleukin 7  
RA receptor.";  
RA J. Exp. Med. 192:659-670 (2000).  
[5] SEQUENCE FROM N.A. (ISOFORM 1).  
STRAIN=FVB/N;  
RA MEDLINE=22386257; PubMed=12477932;  
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
RA Dianchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyay S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalusz D.B.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RA human and mouse cDNA sequences.";  
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[6] SEQUENCE OF 234-359 FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;  
RA MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;  
RA RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -I- FUNCTION: Receptor for thymic stromal lymphopoietin (TSLP). Forms  
CC a functional complex with TSLP and IL7R which is capable of  
CC stimulating cell proliferation through activation of STAT3 and  
CC STAT5. Also activates JAK2. Implicated in the development of the  
CC hematopoietic system.  
CC -I- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1 and  
CC isoform 3). Secreted (isoform 2).  
CC -I- ALTERNATIVE PRODUCTS;  
CC Event-Alternative splicing; Named isoforms=3;  
CC Comment=Additional isoforms seem to exist,  
CC Name=1;  
CC IsoId=Q8CII9-1; Sequence=Displayed;  
CC Name=2; Synonyms=Soluble CRM-2;  
CC IsoId=Q8CII9-2; Sequence=VSP\_008788, VSP\_008789;  
CC Name=3;  
CC IsoId=Q8CII9-3; Sequence=VSP\_008790;  
CC Note=No experimental confirmation available;  
CC -I- TISSUE SPECIFICITY: High level of expression in liver, lung and  
CC testis. Also expressed in heart, brain, spleen, thymus and bone  
CC marrow. Highly expressed in progenitors and myeloid cells. Isoform  
CC 2 is expressed in primary hemopoietic cells.  
CC -I- INDUCTION: Upregulated in the myeloid 32D cell line by granulocyte  
CC colony stimulating factor (G-CSF).  
CC -I- DOMAIN: The WSXWS motif appears to be necessary for proper protein  
CC folding and thereby efficient intracellular transport and cell-  
CC surface receptor binding.  
CC -I- DOMAIN: The BOX 1 motif is important for association with JAKs.  
CC -I- SIMILARITY: Belongs to the type I cytokine family of receptors.  
CC -I- Subfamily 5.  
CC -I- SIMILARITY: Contains 1 fibronectin type III domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; AB039945; BAA92684.1; -  
DR EMBL; AB031333; BAA92159.1; -  
DR EMBL; AF222936; AAF81676.1; -  
DR EMBL; AF201963; AAF82189.1; -  
DR EMBL; BC023788; AAH23788.1; -  
DR EMBL; AK010291; BAB26827.1; -  
DR MGD; MG1:1889506; Tslprt.  
DR InterPro; IPRO03961; FN\_III.  
DR Pfam; PF00041; Fn3; 1  
DR SMART; SM00050; FN3; 1  
DR PROSITE; PS01355; HBWATOPO\_REC\_S\_F1; FALSE NEG.  
KW Receptor; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
FT SIGNAL 1 19  
FT CHAIN 20 359 CYTOKINE RECEPTOR-LIKE FACTOR 2.  
FT DOMAIN 20 232 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 233 253 POTENTIAL.  
FT DOMAIN 254 359 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 117 208 FIBRONECTIN TYPE-III.  
FT SITE 201 205 WSXWS motif.  
FT SITE 262 269 BOX 1.  
FT DISULFD 68 82 BY SIMILARITY.  
FT CARBOHYD 53 53 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 122 122 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT VARSEPLIC 217 217 A -> G (in isoform 2).  
FT FTID-VSP\_008788.  
FT Missing (in isoform 2).  
FT FTID-VSP\_008789.  
FT A -> AGPDCAHLPP (in isoform 3).  
FT VARSPLIC 218 359  
FT VARSPLIC 217 217

Qy

```

OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.; and multiple drug resistant Salmonella
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Incises the DNA at the 3' side of a lesion during
CC nucleotide excision repair. Incises the DNA farther away from the
CC lesion than uvrC. Not able to incise the 5' site of a lesion. When
CC a lesion remains because uvrC is not able to induce the 3'
CC incision, cho incises the DNA. Then uvrC makes the 5' incision.
CC The combined action of cho and uvrC broadens the substrate range
CC of nucleotide excision repair (BY similarity).
CC -!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB008757; AAL20234.1; ALT_INIT.
CC DR StyGene; SG7777; cho.
CC DR InterPro; IPR000305; UvrC N.
CC DR Pfam; PF01541; Exci_endo_N; 1.
CC DR SMART; SM00465; GYGC_1.
CC DR PROSITE; PS50164; UVRC_1.
CC DR SOS response; Hydrolase; Excision nuclease; DNA repair;
CC KW DNA recombination; DNA excision; Complete proteome.
CC SQ SEQUENCE 293 AA; 33332 MW; 4985E05E856F7EE0 CRC64;
CC -----
Query Match 25.4%; Score 54; DB 1; Length 293;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 10 LSAQLLQARLMKEESPVWSWRL 31
DB 89 MGALLLEARLKEQPLFNKRL 110
RESULT 12
CHO_SALTY STANDARD; PRT; 293 AA.
AC Q8ZPU6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog
DE protein).
DE CHO OR STM1309.
OS Salmonella typhimurium.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

```

OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -!- FUNCTION: Incises the DNA at the 3' side of a lesion during
CC nucleotide excision repair. Incises the DNA farther away from the
CC lesion than uvrC. Not able to incise the 5' site of a lesion. When
CC a lesion remains because uvrC is not able to induce the 3'
CC incision, cho incises the DNA. Then uvrC makes the 5' incision.
CC The combined action of cho and uvrC broadens the substrate range
CC of nucleotide excision repair (BY similarity).
CC -!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB008757; AAL20234.1; ALT_INIT.
CC DR StyGene; SG7777; cho.
CC DR InterPro; IPR000305; UvrC N.
CC DR Pfam; PF01541; Exci_endo_N; 1.
CC DR SMART; SM00465; GYGC_1.
CC DR PROSITE; PS50164; UVRC_1.
CC DR SOS response; Hydrolase; Excision nuclease; DNA repair;
CC KW DNA recombination; DNA excision; Complete proteome.
CC SQ SEQUENCE 293 AA; 33362 MW; F422A222A22A05656 CRC64;
CC -----
Query Match 25.4%; Score 54; DB 1; Length 293;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 10 LSAQLLQARLMKEESPVWSWRL 31
DB 89 MGALLLEARLKEQPLFNKRL 110
RESULT 12
CHO_ECO57 STANDARD; PRT; 295 AA.
AC Q8XND7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog
DE protein).
DE CHO OR Z2771 OR ECS2447.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

```

```

RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=1158796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
Ran C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Incises the DNA at the 3' side of a lesion during
CC nucleotide excision repair. Incises the DNA farther away from the
CC lesion than uvrC. Not able to incise the 5' site of a lesion. When
CC a lesion remains because uvrC is not able to induce the 3'
CC incision, cho incises the DNA. Then uvrC makes the 5' incision.
CC The combined action of cho and uvrC broadens the substrate range
CC of nucleotide excision repair (By similarity).
CC -!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005397; BAB56727.1; -;
DR EMBL; AP002558; BAB35870.1; -;
DR PIR; C85783; C85783.
DR PIR; G90934; G90934.
DR InterPro; IPR000305; UvrC.N.
DR SMART; PF01541; Exci_endo_N; 1.
DR PROSITE; PS00465; GIYC; 1.
DR SOS response; Hydrolase; Excision nuclease; DNA repair;
KW DNA recombination; DNA excision; Complete proteome.
KW SEQUENCE 295 AA; 33724 MW; 7CBD7684EEF4321 CRC64;
Query Match 25.4%; Score 54; DB 1; Length 295;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 10 LSAQLLQARLMKESPPVSVRL 31
DB :|||:||||:|:|:|
RESULT 13
CHO_ECOLI
ID CHO_ECOLI STANDARD; PRT; 295 AA.
AC PS9439;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog
DE protein)
GN CHO OR C2140.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H3 / CFT073 / ATCC 700928;
RX MEDLINE=22386234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.;

```

```

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [1]
RP FUNCTION: Incises the DNA at the 3' side of a lesion during
RP nucleotide excision repair. Incises the DNA farther away from the
RP lesion than uvrC. Not able to incise the 5' site of a lesion. When
RP a lesion remains because uvrC is not able to induce the 3'
RP incision, cho incises the DNA. Then uvrC makes the 5' incision.
RP The combined action of cho and uvrC broadens the substrate range
RP of nucleotide excision repair (By similarity).
RP -!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.
RP -----
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation
RP at the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
RP -----
DR EMBL; AE016761; AAN80599.1; -;
DR InterPro; IPR000305; UvrC.N.
DR SMART; PF01541; Exci_endo_N; 1.
DR SMART; SM00465; GIYC; 1.
DR PROSITE; PS00465; UVR_C_1.
DR SOS response; Hydrolase; Excision nuclease; DNA repair;
KW DNA recombination; DNA excision; Complete proteome.
KW SEQUENCE 295 AA; 33694 MW; 9278739AB3E9130E CRC64;
Query Match 25.4%; Score 54; DB 1; Length 295;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 10 LSAQLLQARLMKESPPVSVRL 31
DB :|||:||||:|:|:|
ID CHO_ECOLI STANDARD; PRT; 295 AA.
AC P76213.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog
DE protein)
GN CHO OR B1741.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [2]
RP FUNCTION.
RX MEDLINE=21819418; PubMed=11818552;
RA Moolenaar G.F., van Rossum-Fikkert S., van Kesteren M., Goosen N.;
RT "Cho, a second endonuclease involved in Escherichia coli nucleotide
RT excision repair.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:1467-1472 (2002).
CC -!- FUNCTION: Incises the DNA at the 3' side of a lesion during
CC nucleotide excision repair. Incises the DNA farther away from the
CC lesion than uvrC. Not able to incise the 5' site of a lesion. In

```

vitro, the incision activity of cho is uvrA and uvrB dependent.  
 When a lesion remains because uvrC is not able to induce the 3'  
 incision, cho incises the DNA. Then uvrC makes the 5' incision.  
 The combined action of cho and uvrC broadens the substrate range  
 of nucleotide excision repair.  
 -I- INDUCTION: Repressed by lexA.  
 -I- SIMILARITY: TO THE N-TERMINAL PART OF THE UVR-C FAMILY.

CC	EMBL; AE000269; AAC74811.1; -.
DR	PIR; E64933; E64933.
DR	EcoGene; EGI3933; cho.
DR	InterPro; IPR000305; UvrC.N.
DR	Fram; Ff01541; Excl_endo_N; 1.
DR	SMART; SM00465; Glyc; 1.
DR	PROSITE; PS50164; UVR_C1; 1.
KW	SOS response; Hydrolase; Excision nuclease; DNA repair;
KW	DNA recombination; DNA excision; Complete proteome.
SC	SEQUENCE 295 AA; 33739 MW; DCEADD895F4327 CRC64;

```
Query Match      25.4%; Score 54; DB 1; Length 295;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 11: Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```

QY      10 LSAQLQARLMKEESPVUSWRL 31
      : - - - - - : - - - - -
DQ      89 IGALLLEARLIKEOPLFNKRL 110

```

RESULT 15	
CHO_SHIFL	
ID _CHO_SHIFL	STANDARD; PRT; 295 AA.
AC	P59361;
DT	28-FEB-2003 (Rel. 41, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (UvrC homolog protein).
DE	GENE
DE	CHO OR SF1485 OR SJ602.
OS	Shigella flexneri.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Shigella.
OX	NCBI_TaxID=623;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=301 / Serotype 2a;
RX	MEDLINE=2272406; PubMed=12384590;
RA	Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA	Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA	Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA	Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA	Yu J.
RT	"Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT	through comparison with genomes of Escherichia coli K12 and O157.";
RL	Nucleic Acids Res. 30:4432-4441(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX	MEDLINE=22592074; PubMed=12704152;
RA	Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA	Pournier G., Mayhew G.P., Plunkett G. III, Rose D.J., Darling A.,
RA	Mau B., Ferna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA	Schwartz D.C., Blattner F.R.;
RT	"Complete genome sequence and comparative genomics of Shigella
RT	flexneri serotype 2a strain 2457T.";
RL	Infect. Immun. 71:2775-2786(2003).

CC -1- FUNCTION: Incises the DNA at the 3' side of a lesion during  
CC nucleotide excision repair. Incises the DNA father away from the  
CC lesion than uvrC. Not able to incise the 5' site of a lesion. When  
CC a lesion remains because uvrC is not able to induce the 3'  
CC incision, cho incises the DNA. Then uvrC makes the 5' incision.  
CC The combined action of cho and uvrC broadens the substrate range  
CC of nucleotide excision repair (By similarity).  
CC -1- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC DR EMBL; AEO15171; AAN43077.1; ALT\_INIT.  
CC DR EMBL; AEO16983; AAP16970.1; -.  
CC DR InterPro; IPR000305; UvrC N.  
CC DR Pfam; PF01541; Exci\_endo\_N; 1.  
CC DR SMART; SMO0465; GYC; 1.  
CC DR PROSITE; PSS0164; UVR\_C; 1.  
CC DR SOS response; Hydrolase; Excision nuclease; DNA repair;  
CC KW DNA recombination; DNA excision; Complete proteome.  
CC SQ SEQUENCE 295 AA; 33705 MW; 875D7D594EF4472 CRC64;  
  
Query Match 25.4%; Score 54; DB 1; Length 295;  
Best Local Similarity 50.0%; Pred. No. 6;  
Matches 11; Conservative 5; Mismatches 0; Gaps 0;

QY 10 LSAQLLQARLMKEESPVVSWRL 31  
: : : : : : : :  
Db 89 IGALLLEARLIKEOQPLENKRL 110

Search completed: July 7, 2004, 17:08:12  
Job time : 18 secs

This Page Blank (uspto)